الإحصاء التطبيقي

على الحاسوب

الدكتور
سالم قاسم النعيمي

STATISTICS



الإحصاء التطبيقي
على الحاســــــوب

الطبعة الأولى

1426هـ / 2005م

رقم الإيداع: 2005/5/1148

رقم الإجازة: 2005/5/1108

ردمك: ISBN 9957 – 02 – 188 – 5

دار مجدلاوي للنشر والتوزيع
تليفاكس: ٥٣٤٩٤٩٧ - ٥٣٤٩٤٩٩
ص . ب ١٧٥٨ الجبيهة ١١٩٤١
عمان - الاردن

Dar Majdalawi Pub.& Dis.
Telefax: 5349497 - 5349499
P.O.Box: 1758 Aljubaiha
11941 Amman- Jordan

www.majdalawibooks.com
E-mail: customer@majdalawibooks.com

◆ **الآراء الواردة في هذا الكتاب لا تعبر بالضرورة عن وجهة نظر الدار الناشره.**

تنفيذ وطباعة الرجي بيروت ـ لبنان ـ تلفاكس: ٢٧٢٢٢٥ ٠٠٩٦١١ خليوي: ٣٣٤٦٤٨-٣١٢١٢٤ / ٠٠٩٦١٣

الإحصاء التطبيقي
على الحاســــــــوب

تأليف

د. سالم قاسم حسني السيخ النعيمي

دار مجدلاوي



بسم الـلـه الرحمن الرحيم

﴿ قل هو الـلـه أحد الـلـه الصمد لم يلد ولم يولد

ولم يكن له كفوا أحد﴾ (سور الإخلاص)

صدق الـلـه العظيم

المحتويات

الوحدة الرابعة

مقاييس النزعة المركزية

الوحدة الخامسة

مقاييس التشتت

الوحدة السادسة

الأرقام القياسية

الوحدة السابعة

الارتباط

الوحدة الثامنة

تطبيقات الارتباط والانحدار على الحاسوب

الوحدة التاسعة

الانحدار الخطي

الوحدة العاشرة

الارتباط غير الخطي

الوحدة الحادية عشر

السلاسل الزمنية

المقدمة

الحمد لله رب العالمين، والصلاة والسلام على أشرف الخلق أجمعين سيدنا ومولانا محمد المصطفى وعلى آله وأصحابه أجمعين، وبعد:

إن علم الإحصاء، من العلوم الجوهرية للنهوض بالشعوب إلى الأفضل، وأن استخدام الطرق والأساليب العلمية للإحصاء، تعتبر بمثابة الأدوات، والآلات لصناع القرار. ولقد أصبحت المعلومة الإحصائية البسيطة، تعني عن كثير من السرد والإسهاب، وتوفر الوقت، والجهد، والمال، والموارد للاستفادة منها في نشاط، وعمليات أخرى بحاجة ماسة إليها.

وبدخول الإنسان العربي عصر الكمبيوتر، أصبح علم الإحصاء في متناول يد الجميع، وأصبح العمل الرياضي الشاق، الذي كان ضرورياً لكل صاحب قضية إحصائية، في ذاكرة الكمبيوتر، يطلبه حينما يريد، ونتيجة لذلك، فقد ازدادت الحاجة اليومية في كل مجالات الحياة إلى فهم علم الإحصاء، وتطبيقاته المختلفة، وفي هذا المؤلف سيجد القارئ الكريم الإحصاء التقليدي بالطرق اليدوية، وبنفس الوقت الإحصاء التطبيقي باستخدام الحاسوب –وبلغة البيسك-، وهذا هو ميزة هذا المؤلف، الذي يجمع اليدوي مع الحاسوب في آن واحد، لحل نفس المشكلة، حيث نجد أن المكتبة العربية تفتقر نسبياً إلى مؤلف جامع، رغم غزارة المكتبة العربية بكل المؤلفات في العلوم الرئيسة، بل وفي أفرع هذه العلوم. لهذا وجد الباحث من المناسب تقديم هذا المؤلف لإثراء المكتبة العربية، ولتيسير الحصول على المعلومات والحلول، وعدم التشتت من كتاب إحصاء إلى كتاب حاسوب.

وأملنا أن يملأ هذا المؤلف فراغاً نحن في أشد الحاجة إلى أن يكون مملوءاً لمساعدة ومساندة طلبتنا في المعاهد والجامعات العربية، والباحثين، والمهتمين بعلم الإحصاء والحاسوب وطلبة الدراسات العليا، والجمهور الكريم.

ويحتوي هذا الكتاب على إحدى عشرة وحدة، تتضمن أمثلة وأسئلة، وحلولها النموذجية اليدوية أو الذهنية، والحاسوبية، الذي هو الآن في متناول الجميع.

ففي الوحدة الأولى، تم التطرق إلى أساسيات الإحصاء، وفي الوحدة الثانية تم تناول تبويب البيانات والتوزيعات التكرارية، وتضمنت الوحدة الثالثة التطبيقات الإحصائية بلغة البيسك، وناقشت الوحدة الرابعة مقاييس النزعة المركزية، وتطبيقاتها على الحاسوب، وتناولت الوحدة الخامسة، مقاييس التشتت، وتطبيقاتها على الحاسوب، والعزوم وتطبيقاتها الحاسوبية، وتطرقت الوحدة السادسة إلى الأرقام القياسية، وفي الوحدة السابعة إلى الارتباط، وفي الوحدة الثامنة إلى تطبيقات الارتباط على الحاسوب، وتناولت الوحدة التاسعة، الانحدار الخطي، ومعامل التحديد، والتطبيقات على الحاسوب، وتناولت الوحدة العاشرة، الارتباد غير الخطي، وتمت المناقشة والتعرض إلى السلاسل الزمنية في الوحدة الحادية عشر.

ونسأل اللـه العلي القدير، أن ينتفع بهذا المؤلف المتواضع، وأن يستفيد منه الدارس، والقارئ، وكل من يهتم بالعلم، ويريد الاستسقاء منه.

و اللـه من وراء القصد،

و اللـه ولي التوفيق

المؤلف

الوحدة الأولى

أساسيات الإحصـــاء

1/1 مفهوم علم الإحصاء: **Statistic**

هو العلم الـذي يقـوم بالبحـث في أسـاليب جمـع البيانـات، ووسـائل تحليـل هـذه البيانات للوصول إلى معرفة الظاهرة محل الدراسة.

2/1 خطوات البحث العلمي:

أهم خطوات البحث الإحصائي هي:

1- تحديد الهدف الذي يرمي إليه البحث.
2- تحديد المجتمع المراد دراسته.
3- تحديد المصادر التي تستقى منها البيانات وجمع هذه البيانات.
   أ- المصادر التاريخية.
   ب- المصادر الميدانية.
4- تصنيف البيانات وعرضها.
5- تحليل البيانات.
6- استخلاص النتائج وتفسيرها واتخاذ القرارات المناسبة لحل المشكلة موضع البحث.

3/1 جمع البيانات: **Data Collection**

أنواع الاستمارات الإحصائية:

1- كشف البحث.
2- صحيفة الاستبيان.

- **كشف البحث:** هو كشف يقوم الباحث أو المكلف بجمع البيانـات بملئِـه بنفسـه، تستخدم في حالة انتشار الأمية.

ويعاب عليها بأنه قد يخضع لخطأ التحيز، حيث يمكن للباحث أو المكلف بجمع البيانات، أن يؤثر في إجابات المبحوثين بدون قصد.

- **أما صحيفة الاستبيان:** يقوم المبحوث بملئها.

4/1 أساليب جمع البيانات Data Collection Method

1/1/1 أسلوب الحصر الشامل:

يعتبر هذا الأسلوب على جمع البيانات عن مفردات المجتمع مفردة مفردة، ويتم استخدام هذا الأسلوب في الحالات التالية:

1- إذا كان الغرض من البحث جمع بيانات عن مفردات المجتمع بصفة شخصية شخصية أو فردية.
2- إذا كان الباحث يحتاج إلى دراسة مجتمع معين في مدينة معينة ولا يوجد تقسيمات وكشوف يمكن سحب عينة بطريقة سليمة.
3- إذا كان الباحث يريد الحصول على نتائج على مستوى عال من الدقة. مثل شركات الأدوية، شركات صناعة أنابيب الغاز.
4- إذا كانت مفردات المجتمع المراد دراسته غير متجانسة، وإذا كان المجتمع صغير نسبياً.

2/4/1 أسلوب العينات Sample Method

يعتمد هذا الأسلوب على جمع البيانات من مجموعة مختارة من مفردات المجتمع المراد دراسته، ثم دراسة صفات هذه المجموعة التي اختيرت، ويتم تعميم النتائج التي يحصل عليها الباحث بالنسبة للمجتمع ككل.

مثل دراسة حصر القوة العاملة في الدولة، بحث ميزانية الأسرة.

لماذا نستخدم أسلوب العينات:

1- لأن أسلوب العينات يوفر الوقت والجهد والتكاليف اللازمة لإجراء البحث.

2- لأن أسلوب العينات يؤدي على الإقلال من مدى التحيز الناتج عن عدم الدقة في القياس وذلك لمحدودية مفردات المجتمع المراد دراسته.

3- عندما يكون المجتمع لا نهائي.

4- عندما أسلوب الحصر الشامل يغني المفردات محل الدراسة، كما في حالة فحص دم المريض.

هناك أخطاء يتعرض لها الباحث عند استخدام أسلوب العينات.. وهناك نوعان من الخطأ العشوائي:

1- خطأ الصدفة . Random Error

2- خطأ التحيز. Bias Error

خطأ الصدفة:

أ- عدم التجانس في مفردات المجتمع. فكلما كانت مفردات المجتمع غير متجانسة، كلما زاد احتما تعرض الباحث لخطأ الصدفة.

ب- حجم العينة المسحوبة بالنسبة لحجم المجتمع الذي سحبت منه هذه العينة فكلما كان حجم العينة بالنسبة لحجم المجتمع الذي سحبت منه كبيراً كلما قل احتمال تعرض الباحث لخطا الصدفة.

أما خطأ التحيز:

فهو ذلك الخطا الذي ينشأ نتيجة لعوامل إنسانية بحتة، ويحدث خطا التحيز للأسباب التالية:

سوء اختيار العينة، أي أن العينة لا تمثل المجتمع الذي سحبت منه، ويحدث عندما يتم الاختيار على أساس شخصي وخطا التحيز يشكل خطراً كبيراً على نتائج العينة لصعوبة تقديره.

ويمكن تلافي خطا التحيز يمكن تلافيه بواسطة قوانين الاحتمالات وبالتالي الحصول على نتائج مقبولة.

1/4/1 أنواع العينات

هناك عدة أنواع للعينات، والاختيار لطريقة معينة يبنى على اعتبارات معينـة مثـل: طبيعة التباين، والاختلاف بين مفردات المجتمع المراد دراسـته، والتكـاليف التـي يتحملهـا الباحث.

وبأخذ هذه الاعتبارات بعين الاعتبار، يمكن للباحـث اختيـار طريقـة اختيـار العينـة العشوائية.

وأهم أنواع العينات:

1/3/4/1 العينة العشوائية البسيطة Simple Random Sample

ويـتم الاختيـار في هـذا النـوع مـن العينـات بإعطـاء فـرص متكافئـة لكـل مفـردات المجتمـع عنـد الاختيـار، أي أننـا نعطـي لكـل مفـردة مـن مفـردات المجتمـع نفس الفرصة من الاختيار.

وتمتاز العينة العشوائية البسيطة بسهولة اختيارها وبساطتها.

2/3/4/1 العينة الطبقية:

يـتم تقسـيم المجتمـع المـراد دراسـته إلى طبقـات متجانسـة مـن ظـاهرة لهـا علاقـة بـالمتغير المطلـوب دراسته، ويـتم بعـد ذلـك سـحب عينـة عشـوائية بسـيطة مـن كل طبقـة عـلى حـده. بحيـث تكـون نسـبة مفـردات الطبقـة في العينـة مسـاوية لنسـبة مفردات الطبقة في المجتمع ككل.

3/3/4/1 العينة المرحلية: Multiple Stage Sample

يـتم اختيارهـا بتقسـيم المجتمـع إلى مجموعـات، ثـم سـحب عينـة عشـوائية مـن المجموعات نفسها.

ويستخدم هـذا النـوع مـن العينـات في الحـالات التـي يكـون فيهـا المجتمـع محـل الدراسة كبيراً جداً.

4/3/4/1 العينة المنتظمة: Systematic Sample

الوحدة الثانية

تبويب البيانـــــات

Data Classification

1/2 التوزيعات التكرارية **Frequency Distributions**

يلجأ الباحث إذا كان عدد القيم كبيراً إلى تقسيم القيم الأوليى إلى مجماعات جزئيـة تسمى كل منها فئة تكرارية.

أي يقسم الباحث القيم الأولية إلى فئات تشمل كل فئة عـدد مـن القـيم الأوليـة أو البيانات الخام المتقاربة من بعضها البعض.

ويحدد الباحث الحد الأدنى والحـد الأعـلى لكـل فئـة ثـم يـوزع القـيم الأوليـة عـلى الفئات بأن يصنع كل مفردة في الفئة المناسبة لها، ثم يعد المفردات الموجودة من كل فئة، ويضع العدد أمام كل فئة، فيحصل على جدول التوزيع التكراري.

فالتوزيع التكراري إذن، عبارة عن توزيع يبين توزيع البيانـات الخـام إلى مجموعـات أو فئات.

مثـــال:

قيست أطوال عينة من 50 مفردة، فكانت نتيجة تلك الأطوال، هي:

168	184	175	182	168	190	162	183	171	166
175	182	160	153	157	168	178	162	149	169
160	172	175	159	175	182	191	173	167	176
166	175	182	184	177	169	174	168	160	199
190	160	179	183	171	179	162	169	197	185

المطلوب:

عرض هذه البيانات في جدول توزيع تكراري؟

من البيانات نرى أن أصغر قيمة هي 149 وأكبر قيمة هي 199.

يمكن تقسيم تلك القيم الأولية إلى 6 ستة فئات تشمل الأولى كل من طولهم من 140 وأقل من 150سم.

والثانية تشمل هؤلاء الذين يقع طولهم من 150 إلى أقل من 160سم، وهكذا إلى أن نصل إلى الفئة الأخيرة.

وقد استخدم هذا النوع من التقسيم لأن التغير المدروس (الطول) متغير مستمر، والجدول التالي يبين ذلك.

الفئة	التفريغ	التكرار
140 وأقل من 150	1	1
150 وأقل من 160	111	3
160 وأقل من 170	11 ~~1111~~ ~~1111~~ ~~1111~~	17
170 وأقل من 180	~~1111~~ ~~1111~~ ~~1111~~	15
180 وأقل من 190	~~1111~~ ~~1111~~	10
190 وأقل من 200	1111	4
		50

وفي الصورة النهائية يلغى الجزء الخاص بتفريغ البيانات الأولية، ويكون الجدول مكوناً من عمودين فقط، يوضح أحدهما الفئات، والثاني التكرارات الخاصة بكل فئة، ويسمى هذا بالجدول التكراري البسيط.

فالتوزيع التكراري، عبارة عن جدول يتم فيه توزيع المشاهدات المأخوذة عن ظاهرة معينة على عدد معين من الفئات تحدده ظروف المسألة.

وتتلخص أهم خطوات بناء جدول توزيع تكراري لأي عدد من البيانات في الآتي:

1- تحديد عدد الفئات.

2- تحديد طول الفئة.

3- إيجاد جدول الفئة.

4- إيجاد عدد التكرارات في كل فئة.

نجد طول الفئة بإيجاد المدى العام، ثم تقسيمه على عدد الفئات:

فلو كانت لدينا أكبر قمية 199 كما في المثال، وأقل قيمة هي 149

فإن المدى العام = 199 – 149 = 50

التمثيل البياني للتوزيع التكراري:

يفيد في إظهار الخصائص الرئسيسة للتوزيع مما يسهل متابعة التغير في الظاهرة.

ويختلف التمثيل البياني للتوزيعات التكرارية، حسب توزيع الفئات.

التوزيعات التكرارية ذات الفئات المتساوية، تختلف طريقة تمثيلها بيانياً عن التوزيعات التكرارية ذات الفئات غير المتساوية، وإن كان الأساس النظري للتمثيل البياني واحد في الحالتين.

وهناك طريقة عديدة لتمثيل التوزيعات التكرارية، منها:

1- المدرج التكراري.
2- المضلع التكراري.
3- المنحنى التكراري.

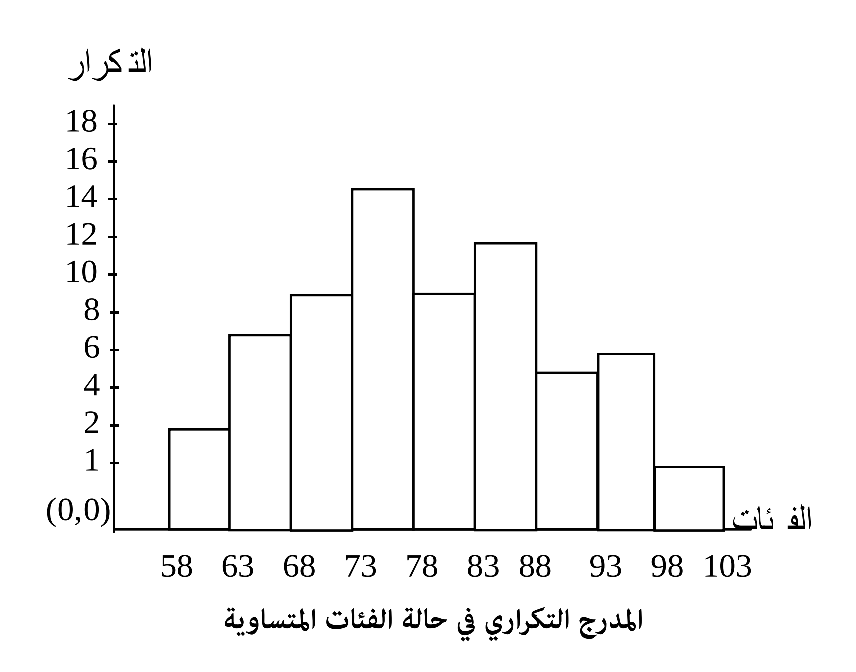
1/1/2 المدرج التكراري

1/1/1/2 في حالة الفئات المتساوية:

المدرج التكراري عبارة عن شكل بياني، يحتوي على مجموعة من المستطيلات المتلاحقة، تتناسب مساحتها مع تكرارات الفئات. أي أنه في هذا النوع يمثل التكرار بمساحة مستطيل، وأن المحور الأفقي يمثل الفئات. وأن قواعد المستطيلات ستكون متساوية. وفي حالة (الفئات المتساوية)، فإن ارتفاع كل من المستطيلات يمثل تكراراً لفئة معينة، ويمكن بيان ذلك، بالمثال التالي:

ارسم المدرج التكراري للتوتر مع الآتي:

الفئـــة	التكرار
58- 63	2
64-	6
68-	8
73-	15
78 -	10
83 -	12
88-	5
93 -	6
98 - 103	1



المدرج التكراري في حالة الفئات المتساوية

2/1/1/2 في حالة الفئات غير المتساوية:

تكون القاعدة مختلفة لكل مستطيل، ولا بد من حساب ارتفاع المستطيلات، بحيث تتناسب المساحات مع قيمة التكرارات، ويتم ذلك عن طريق تعديل تكرار كل فئة طبقاً لطول الفئة، فنحصل على تكرارات معدل لأغراض الرسم البياني.

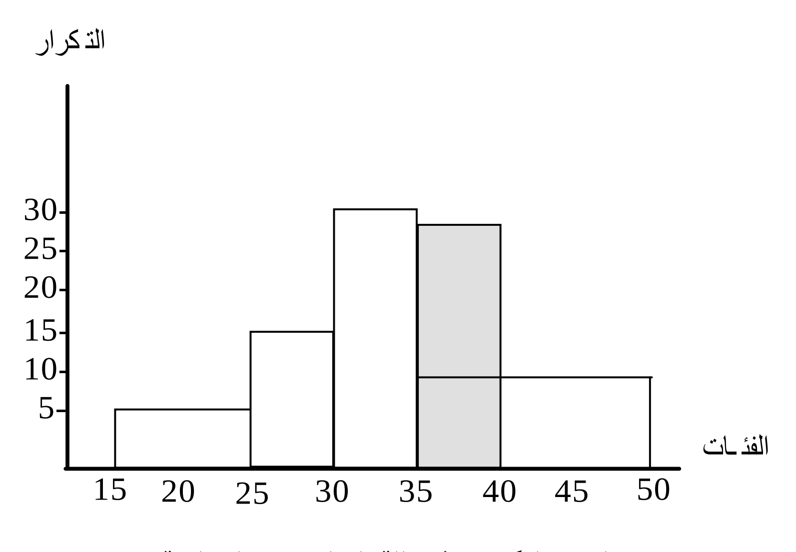
مثال:

التكرار	الفئـــة
40	15 -
80	25 -
150	30 -
140	35 -
90	40 - 50

وحيث أن فئات هذا التوزيع التكراري غير متساوية، فإنه يجب تعديل التكرارات، بقسمة كل تكرار على ÷ طول لفئة قبل الرسم، كالآتي:

التكرار المعدل	التكرار الأصلي	طول الفئة	الفئة
4	40	10	15 -
16	80	5	25 -
30	150	5	30 -
28	140	5	35 -
9	90	10	40 - 50

ثم نرسم على المحور الرأسي التكرارات المعدلة، وعلى المحور الأفقي الحدود الدنيا للفئات، ونقيم أعمدة يمثل طولها التكرار المعدل، ثم نوصلها بخطوط أفقية، فنحصل على المدرج التكراري.



المدرج التكراري في حالة الفئات غير المتساوية

2/1/2 المضلع التكراري

1/2/1/2 في حالة الفئات المتساوية:

المضلع التكراري عبارة عن خط نقاط تمثل كل منها تكرار فئة معينة ثم رصده أمـام مركزه، مع غلق المضلع من طرفيه باستخدام مراكز الفئات قبل الأولى وبعد الأخيرة، حيث أن تكرار هاتين الفئتين صفرًا.

ويمكن رسم المضـلع التكـراري مبـاشرة مـن التوزيـع التكـراري أو باسـتخدام مـدرج تكراري أولاً، ثم توصيل النقاط التي تقع في منتصف قمم المسـتطيلات، ثـم غلـق المضـلع بإدخال الفئات التي تكرارها يساوي صفر.

مثال:

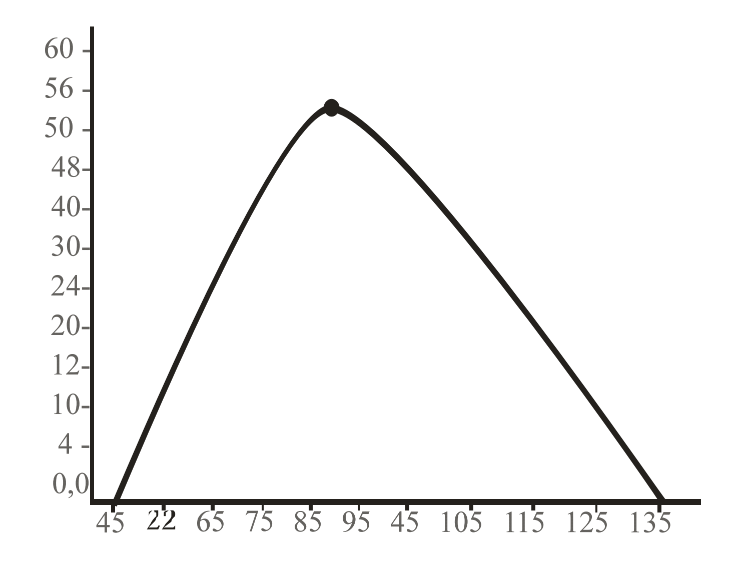
التكرار	الفئة	مراكز الفئات
0	40 -	2 ÷ (50+40)
4	50 -	45
24	60 -	55
48	70 -	65
56	80 -	75
48	90 -	85
40	100 -	95
24	110 -	105
12	120 -	115
	130 -	125
		135

(قبل الأولى)

$\frac{140+130}{2}$

$\frac{140+130}{2}$

(بعد الأخيرة)



2/2/1/2 في حالة الفئات غير المتساوية

لا بد من تعديل التكرارات، وذلك بقسمة ÷ تكرار كل فئة على ÷ طولها أوّلاً، ثم نستخدم التكرارات المعدلة في الرسم البياني.

يخصص المحور الأفقي لمراكز الفئات، والمحور الرأسي للتكرارات المعدلة، ثم نرصد أمام كل الفئات قبل الأولى وبعد الأخيرة ذات التكرار المساوي للصفر، ويمكن تطبيق ذلك على المثال السابق.

3/1/2 المنحنى التكراري

1/3/1/2 في حالة الفئات المتساوية:

المنحنى التكراري عبارة المضلع التكراري بعد تمهيده أي (استبدال الخط البياني المنكسر الممثل للمضلع التكراري، بمنحنى ممهد، يصل بين كل النقاط الممثلة للتكرارات، عند مراكز الفئات، أو يمر باكبر عدد منها.

- يخصص المحور الأفقي لمراكز الفئات، والمحور العمودي للتكرارات.
- يمكن استخدام الحد الأدنى للفئات.
- نضع كل تكرار مقابل لنقطة المركز والمحور الرأسي للتكرارات.
- يمهد بمنحنى يصل بين هذه النقاط.
- تكون المساحة المحصورة بين المحور والمنحنى التكراري، هي مجموع التكرارات.

2/3/1/2 في حالة الفئات غير المتساوية

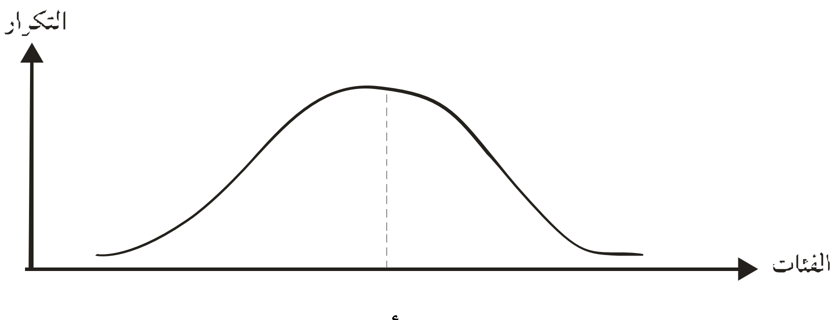
لا بد من اتباع القاعدة السابقة لتعديل التكرارات، وذلك بقسمة ÷ التكرارات الأصلية ÷ على طول الفئة، واستخدام التكرارات المعدلة بدلاً من الأصلية على المحور الرأسي في الرسم البياني.

3/3/1/2 أشكال منحنيات التوزيعات التكرارية

يختلف شكل المنحنى طبقاً لطبيعة التوزيع التكراري تحت الدراسة.

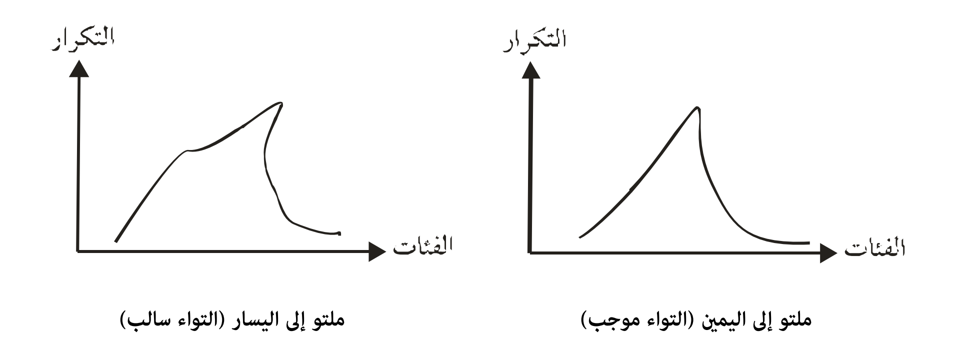
هناك منحنيات متماثلة، والتي تنقسم إلى نصفين متطابقين تماماً، إذا أسقطنا عمـودًا من قمة المنحنى إلى المحور الأفقي، ويأخذ هذا، النوع من المنحنيات شكل الجرس.

وتختلف المنحنيات تفرطحًا أو تدببًا، حسب حجم التكرارات على جانبي القمة، كما يتضح من الشكل التالي:



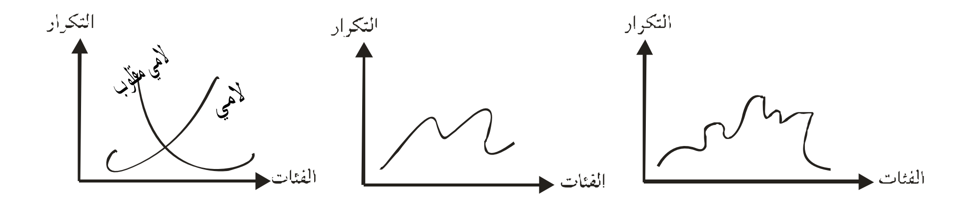
منحنى شكل الجرس أو منحنى متماثل

وهناك منحنيات غير متماثلـة، حيـث تتركـز أغلـب التكـرارت عنـد أصـغر القـيم أو أكبرها، بدلاً من تركزها عند القيمة الوسطى، كما في حالة المنحنى التماثـل ولا تنقسـم إلى قسمين متساويين متطابقين على بعضهما لاختلاف انحدار جانبي المنحنى، لذا تسمى هذه المنحنيات بالمنحنيات الملتوية، أما إلى اليمين (التـواء موجـب) إذا تركـزت التكـرارات عنـد القيم الصغرى، فيصعد المنحنى بسرعة، ويهبط ببطء. أو ملتو إلى اليسـار (التـواء سـالب) عندما تتركز التكرارات عند القيم الكبرى، فنجد ا لمنحنى يصعد ببطء ويهبط بسرعة.



ملتو إلى اليمين (التواء موجب) ملتو إلى اليسار (التواء سالب)

وهناك منحنيات متزايدة أو متناقصة تسمى أحياناً لاميه أو لاميه مقلوبه، وقد يكون للمنحنى قيمتين أو له قمم متعددة، كما يتضح في الأشكال التالية:



4/1/2 التوزيعات التكرارية

1/4/1/2 التوزيعات التكرارية المتجمعة:

تستخدم لمعرفة عدد المفردات التي تكون أعلى أو أقل من قيمة معينة، وتعكس الوضع النسبي لمفرده ما في البيانات الأولية.

- عندما يتم تجميع تكرارات كل فئة مع تكرارات الفئات السابقة لها، نحصل على التوزيع التكراري المتجمع الصاعد.

- ويقيس هذا التوزيع، التكرارات الفئات السابقة لها، نحصل على التوزيع التكراري المتجمع الصاعد.

- ويقيس هذا التوزيع، والتكرارات التي تقل عن الحد الأعلى لكل فئة.

- التكرار المتجمع الصاعد لكل فئة عبارة عن التكرار الأصلي لهذه الفئة + مضافاً إليها تكرارات الفئات السابقة عليه. لذلك يبدأ التكرار المتجمع الصاعد بتكرار الفئة الأولى، حيث لا يوجد فئات سابقة لها ثم نضيف + التكرار المتجمع الصاعد للفئة السابقة إلى تكرار الفئات الأخرى للحصول على التكرار المتجمع الصاعد.

	حدود الفئة	التكرار	التكرار المتجمع الصاعد		
أقل من	140	0	0		
أقل من	150	1	1		
أقل من	160	3	4	=	1 + 3
أقل من	170	17	21	=	4 + 17
أقل من	180	15	36	=	21 + 15
أقل من	190	10	46	=	36 + 10
أقل من	200	4	50	=	46 + 4

يستنتج من أن التكرار المتجمع الصاعد، ينتهي بالعدد الكلي للمشاهدات ومنه نستنتج أن هناك 36 فرداً أطوالهم اقل من 180 سم، وهكذا.

ويمكن إيجاد التوزيع التكراري المتجمع النازل لنفس المثال:

الحد الأدنى للفئة	التكرار المتجمع النازل
140 فأكثر	1 + 49 = 50
150 فأكثر	3 + 46 = 49
160 فأكثر	17 + 29 = 46
170 فأكثر	15 + 14 = 29
180 فأكثر	10 + 4 = 14
190 فأكثر	4

مثال:

التكرار	الفئة
2	1.5 -
1	2.0 -
4	2.5 -
15	3.0 -
10	3.5 -
5	4.0 -
3	4.5 – 5.0

من هذا التوزيع التكراري، احسب أولاً، التوزيع التكراري المتجمع الصاعد ثم النازل.

الفئة	التكرار المتجمع الصاعد	التكرار المتجمع الصاعد النسبي
أقل من 2.0	2	5.0
أقل من 2.5	3	7.5
أقل من 3.0	7	17.5
أقل من 3.5	22	55.0
أقل من 4.0	32	80.0
أقل من 4.5	37	92.5
أقل من 5.0	40	100.0

ويتم هـذا بـالتغيير عـن التكـرارات المطلقـة المقابلـة لكـل فئـة كنسـبة مئويـة مـن المجموع الكلي (في حالة التكرار المتجمع الصاعد هو تكرار الفئة الأخيرة) وهـذا مـا تـم في العمود الثالث من الجدول السابق.

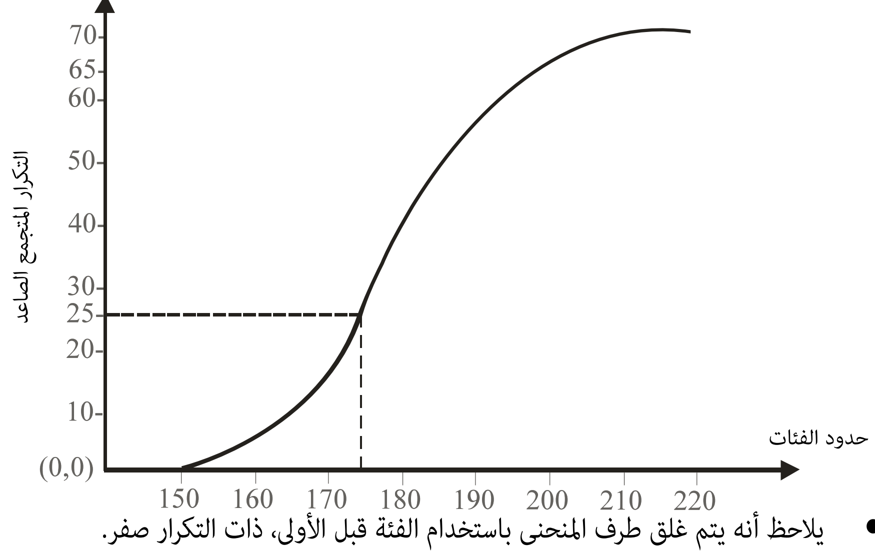
التمثيل البياني للتوزيعات التكرارية المتجمعة

يخصص الإحداثي الأفقي للفئات، والإحداثي الرأسي للتكرارات المتجمعة.

ويمثل التوزيع التكراري المتجمع الصاعد بـالمنحنى المتجمـع الصـاعد، كـما في المثـال التالي:

الحدود العليا للفئات	التكرار المتجمع الصاعد
أقل من 160	8
أقل من 170	18
أقل من 180	34
أقل من 190	48
أقل من 200	58
أقل من 210	63
أقل من 220	65

ترصد النقاط الواردة بالتوزيع التكراري أمام الحدود العليا للفئات.



- يلاحظ أنه يتم غلق طرف المنحنى باستخدام الفئة قبل الأولى، ذات التكرار صفر.
- يفيد هذا المنحنى في معرفة التكرارات المتجمع التي تقل فيها الظاهرة عن (قيمة معينة).

فمثلاً إذا أردنا معرفة التكرارات التي تقل عن 175 (وهي قيمة غير واردة بالجدول).

- فمثلاً إذا أردنا معرفة التكرارات التي تقل عن 175 (وهي قيمة غير واردة بالجـدول الأصلي).

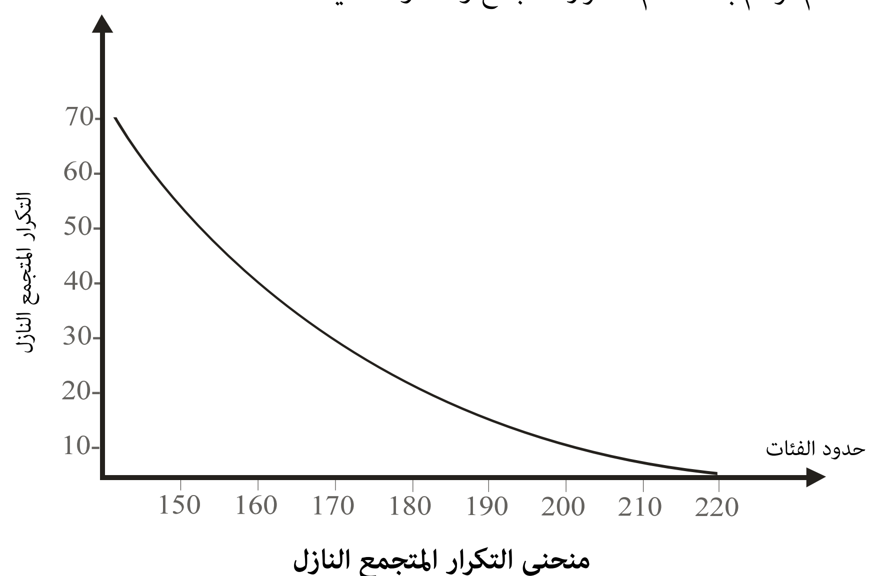
- نقيم عمود من المحور الأفقي عند هذه النقطة إلى أن يلتقي بالمنحنى الصاعد، ونقرأ التكرار عند هذه النقطة وهو 25.

أما تمثيل التوزيع التكراري المتجمع النازل، فيتم باستخدام المنحنى المتجمع النازل.

ويتبع نفس الاسلوب السابق، حيث نستخدم الحدود الدنيا للفئات على المحول الأفقي بدلاً من الحدود العليا.

الحدود الدنيا للفئات	التكرار المتجمع النازل
150 فأكثر	65
160 فأكثر	57
170 فأكثر	47
180 فأكثر	31
190 فأكثر	17
200 فأكثر	06
210 فأكثر	02

ثم نرسم باستخدام التكرار المتجمع والحدود الدنيا للفئات

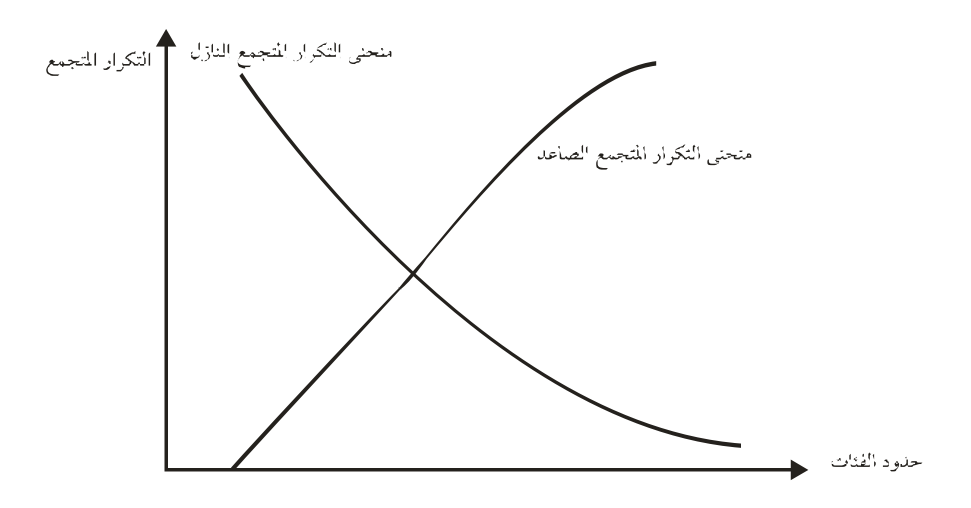


منحنى التكرار المتجمع النازل

ويمكن استخدام نفس الأسلوب السابق في تحديد التكرارات التي تزيد عن قيمة معينة، أو تلك التي تقع داخل فئة غير واردة في الجدول الأصلي.

بإقامة أعمدة من المحور الأفقي إلى المنحنى المتجمع النازل، وقراءة التكرار المقابل.

كما يمكن تمثيل كل من المنحنى المتجمع الصاعد والنازل على رسم بياني واحد، بتخصيص المحور الأفقي للفئات، والمحور الرأسي للتكرارات، ثم نقوم برصد النقاط أمام الحد الأعلى أو الأدنى للفئة حسب نوع المنحنى، كالآتي:



2/4/1/2 التوزيعات التكرارية المزدوجة:

كثيرا ما يحتاج الباحث إلى دراسة العلاقة بين ظاهرتين أو متغيرين، مثل العلاقة بين السعر والكمية المطلوبة، أو بين السعر والكمية المعروضة من سلعة معينة، أو بين المستوى العلمي والمرتب أو بين طول الأفراد وأوزانهم...الخ. لذلك.

لا نستطيع استخدام التوزيعات التكرارية البسيطة، بل لا بد من تبويب البيانات المتعلقة بالمتغيرين ويسمى «جدول تكراري مزدوج مقسم أفقياً طبقاً لفئات إحدى الظاهرتين أو المتغيرين محل الدراسة، ورأسياً طبقاً لفئات الظاهرة أو المتغير الآخر».

يمكن عرض الجدول التالي في توزيع تكراري مزدوج

فئات المتغير	أفقي X	رأسي Y	أفقي X	رأسي Y
	12	14	28	20
	16	27	12	15
	28	28	22	25
	22	33	15	22
	32	44	24	35
	38	52	29	40
	15	16	17	28
	21	29	24	20
	11	24	20	50
	23	44	18	21
	29	35	35	29
	25	26	38	49
			18	32

نركز على كل متغير بمفرده، ونحدد عدد وأطوال الفئات التي تغطى كل المشاهدات المتعلقة بهذا المتغير، ثم تجمع + التقسيمين في جدول واحد، فإذا جعلنا التقسيم الرأسي يمثل فئات المتغير Y الذي سيقسم إلى 5 فئات طول كل منها 10 وتبدأ من 10.

في حين أن التقسيم الأفقي يمثل المتغير x الذي سيقسم إلى ثلاثة فئات فقط تبدأ من 8 وطول كل فئة 10، ثم نأخذ كل المشاهدتين متناظرتين ونفرغها في الخانة الملائمة... كالآتي:

X / Y	8 -	18 -	28-38	مجموع y
10 –	3	5		3
20 –	4	6	2	12
30 –	-	3	1	4
40 –	-	1	3	4
50 –	-	1	1	2
مجموع x	7	11	7	25

يتضح من الجدول أن هناك حالة واحدة تزيد فيها قيمة Y عن 50 وتزيد X المقابله عن 28، في حين توجد ثلاث حالات تقل قيمة Y فيهم عن 20، وتقل قيمة X المناظرة عن 18.

ويمكن ارتباط جداول تكرارية منفصلة من الجدول السابق يمثل الأول توزيع تكراري بسيط للمتغير X، ويمثل الثاني توزيع تكراري بسيط للمتغير Y.

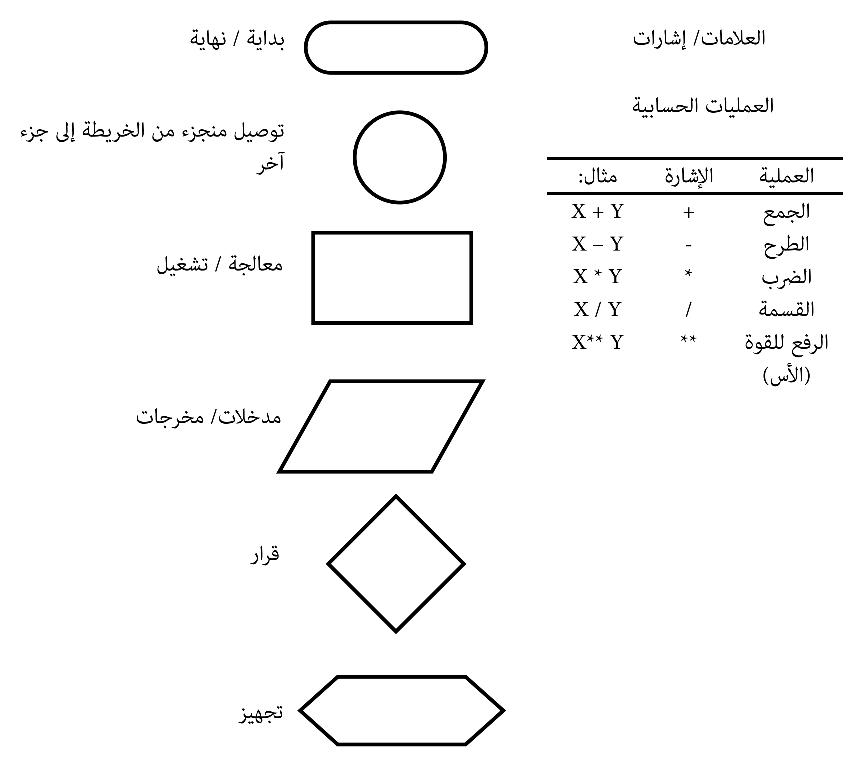
ويستخدم مثل هذا النوع من الجداول في التحليل الإحصائي لدراسة العلاقة بين، عندما تكون المشاهدات المتوفرة عن كل منهما وفيرة العدد، حيث يوفر التبويب المزدوج الكثير من الوقت والجهد.

الوحدة الثالثة

التطبيقات الإحصائية بلغة البيسك

1/3 خريطة سير العمليات **Flow Chart**

الرموز وفق المعهد الأمريكي الوطني للمواصفات والمقاييس.



العلامات/ إشارات

العمليات الحسابية

العملية	الإشارة	مثال:
الجمع	+	X + Y
الطرح	-	X – Y
الضرب	*	X * Y
القسمة	/	X / Y
الرفع للقوة (الأس)	**	X** Y

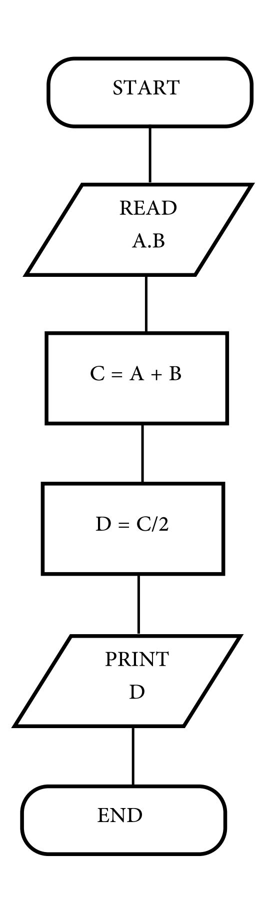
«الأشكال الأساسية المستخدمة في رسم خرائط سير العمليات»

BASIC: Beginners All-Purposes Symbolic Instruction code.

وتعني: اللغة الرمزية المتعددة الأغراض للمبتدئين.

مثال: ارسم خريطة سير عمليات لقراءة رقمين وطباعة الوسط الحسابي لهما.

الحل:



هناك ملاحظات ينبغي ذكرها:

1- كل خرائط سير العمليات تبدأ بـ (بداية START) وتنتهي بـ نهاية (END).

2- انسيبا الخريطة يكون من أعلى إلى أسفل، ما لم تعترضه تحويلات تغير مساره.

3- كل خطوة (شكل) من خطوات الخريطة ينبغي أن تكون متصلة من جانبين لتوضيح الخطوة السابقة، والخطوة التي تليها، (ما عدا بالطبع البداية START والنهاية END.

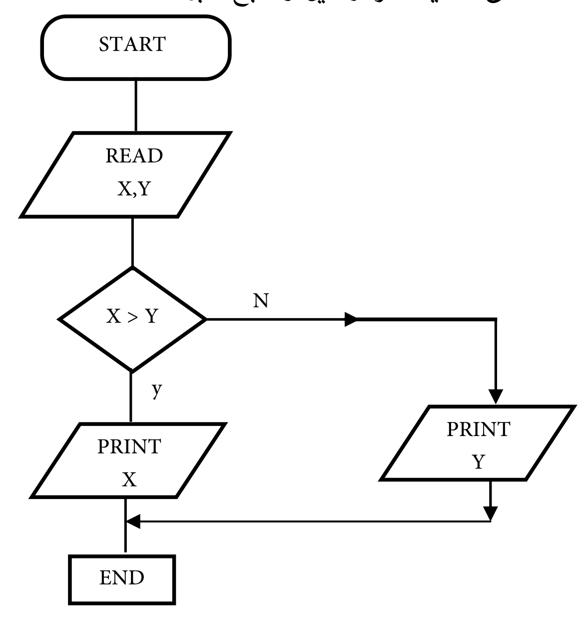
2/3 تحويل المسار: BRANCHING

انسياب خرائط سير العمليات يتحول في مرحلة من المراحل إلى أحد المسارات أو الآخر، اعتماداً على نتيجة قرار معين، لذلك فهو يستخدم شكل القرار DECISION والذي يكون نتيجة نعم أو لا.

وهنا تستخدم الرموز والإشارات التي تختبر العلاقة بين قيمتين RELATIONAL كالآتي:

الإشارة	مدلولها
=	يساوي
>	أكبر من
<	أصغر من
> =	أكبر من أو يساوي
=>	أصغر من أو يساوي
< >	لا يساوي

مثال: خريطة التدفق التالية تقرأ رقمين وتطبع أكبرهما فقط.



3/3 الدوارة/ الدورة LOOP

من خصائص الكمبيوتر، وهي قدرته على تكرار عملية معينة أو مجموعة عمليات، أي عدد من المرات، ويسمى هذا التكرار بالدوارة Loop الدورة.

مثال: اقرأ درجات 30 طالباً في امتحان معين، واحسب متوسط الدرجات وطباعته.

لاحظ الآتي في هذا المثال:

- استخدم رمز التهيئة لتنظيف حقل المجموع T والعداد C وذلك بوضع القيمة صفر فيهما كقيمة ابتدائية، لأن الأول T يستخدم لعملية الجمع التراكمي للدرجات، والثاني لعد الدرجات نفسها حتى تحدد نهاية الدوارة/الدورة Loop.
- استخدم متغير واحد هو X لقراءة كل القيم، وبهذا يأخذ قيمة متغيرة في كل دورة، وفي كل دورة يضاف قيمة X إلى المجموع السابق T ونضيف 1 إلى العداد C.
- لتحديد نهاية الدوارة يستخدم رقم القرار ◇ لمعرفة إذا كانت الأرقام كلها قد قرئت.

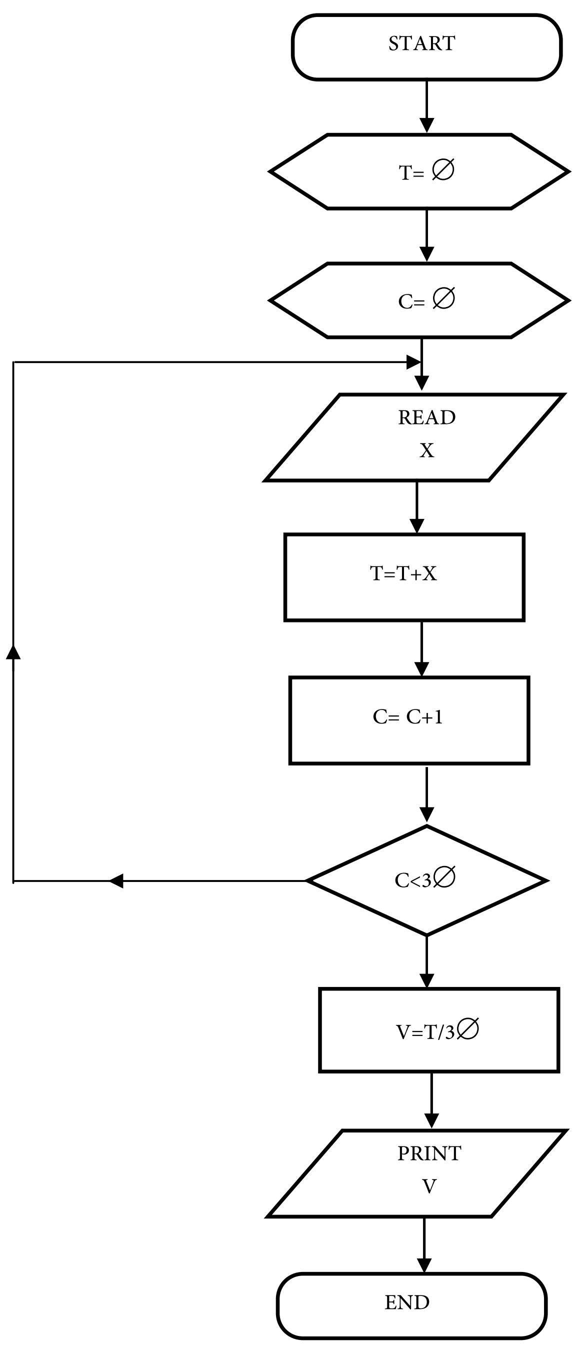
إذ نسأل إذا كان العداد C أقل من 30 وهي قيمته التوائية، فإذا كانت الإجابة بلا، فإننا نرجع لبداية الدوارة، وإلا فإننا نحسب الوسط الحسابي ونطبعه وننهي الخريطة.

أنماط خرائط سير العمليات

إن أي خوارزمية يمكن تمثيلها بإحدى الطرق الثلاث:

1- منطق تسلسلي SEQUENCE
2- منطق اختيار SELECTION
3- منطق تكرار ITERATION

FLOW CHART FOR LOOP

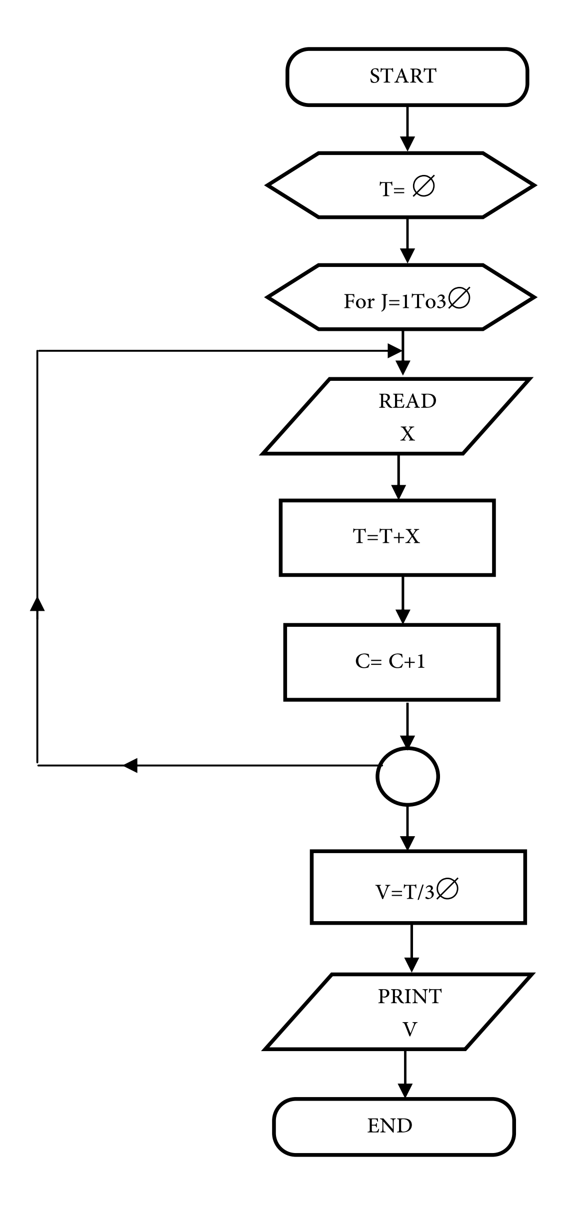


منطق التكرار LOOP:

في لغة البيسك BASIC يأخذ عبارة FOR… NEXT

مثل: FOR J = 1 TO N STEP1

شكل يوضح حل مسألة لقراءة 30 رقماً وطباعة متوسطها.



اكتب برنامج قيمتـين رقميتـين، اجمعهـما، واحسـب متوسـطهما، ثـم اطبـع الـرقمين والمجمـوع والمتوسط:

```
10  REM PROGRAM TO ADD
      TWO NUMBERS
20    REM AND TO PRINT THEIR SUM AND AERAGE
30    X = 52
40    Y= 34
50    S = X+Y
60    V=S/2
70    PRINT X,Y
80    PRINT S,V
90    END
```

عند تنفيذ هذا البرنامج ستظهر المخرجات التالية:

```
52          34
86          43
```

يمكن للسطر رقم 80 أن يكون كالآتي:

```
80    PRING, TOTAL =  '; S;  ,  AVERAGE = ';V
```

وعندها كانت المخرجات ستكون:

```
TOTAL = 86          AVERAGE = 43
```

أوامر الإدخال: لامداد البرنـامج بالبيانـات، نسـتخدمهما READ, INPUT LET عنـد تنفيـذ أي من عبارات INPUT فإنه تظهر على الشاشة علامة استفهام (؟) وما على المسـتخدم حينئـذ إلا إدخـال القيمة المطلوبة.

```
10    INPUT A
20    PRINT
30    END
```

وعند التنفيذ

```
RUN
 ?8
  8
```

4/3 البرامج الفرعية SUBPOUTINES

مثال: البرنامج التالي مكتوب بطريقة البرامج الفرعية، وهو يقوم بقراءة أسماء عشرة طلاب، ودرجات كل منهم في 3 اختبارات، ومن ثم يحسب المجموع والمعدل لكل، ويطبع كل البيانات.

برنامج لقراءة أسماء ودرجات 10 طلاب في 3 اختبارات وحساب مجموع ومعدل كل منهم وطباعتها.

البرنامج: يستخدم طريقة البرامج الفرعية.

```
10    REM
20    REM
25    REM SUBROUTINE
30    DIM N S(10), A(10), B(10), C(10), T(10), V(10)
40    PRINT USING  400
50    PRINT USING  410
55    PRINT
60    GO SUB 100
70    GO SUB 200
80    GO SUB 300
90    STOP
100   FOR I=1 TO 10
130   READ N S(I), A(I), B(I), C(I), C(I) REM      الاسم والدرجات الثلاث
140   NEXT I
150   RETUR  N
200   FOR I = 1 TO 10
210   T(I)= A(I)+B(I)+C(I) REM المجموع
220   V(I) = T (I)/3 REM المعدل
```

```
230     NEXT I
240     RETURN
300     FOR I = 1 TO 10
310     PRINT USING 500, V(I), T(I), C(I), B(I), A(I), A(I), N S (I), I
320     PRINT USING 510
330     NEXT I
340     PETURN
             400     م الاسم درجة 1 درجة 2 درجة 3 مجموع معدل
410     ________________________________________
500      ** ** ***  ***  ***     **    *******   C **
510     ________________________________________
600     DATA   حازم الشيخ    75 72 76    محمد رجب  80 82 85
610     DATA   محمد الرفاعي   90 88 86    زكي محمد  82 91 88
620     DATA   سالم قاسم     81 75 88   طارق الصائغ  65 52 60
630     DATA   عامر يحيى     89 91 93   عدنان محمد  89 91 93
999     END
```

المخرجات:

م	الاسم	درجة 1	درجة 2	درجة 3	مجموع	معدل
1	زكي محمد	82	91	88	261	87.00
2	محمد رجب	80	82	85	247	82.33
3	حازم الشيخ	75	72	76	223	74.33
4	سالم قاسم	81	75	88	244	81.33
5	طارق الصائغ	65	52	60	177	59.00
6	محمد الرفاعي	90	88	86	264	88.00
7	كمال الشيخ	89	91	93	273	91.00
8	عامر يحيى	75	76	80	231	77.00
9	عدنان محمد	85	84	88	257	85.67
10	عصام النعيمي	89	91	93	273	91.00

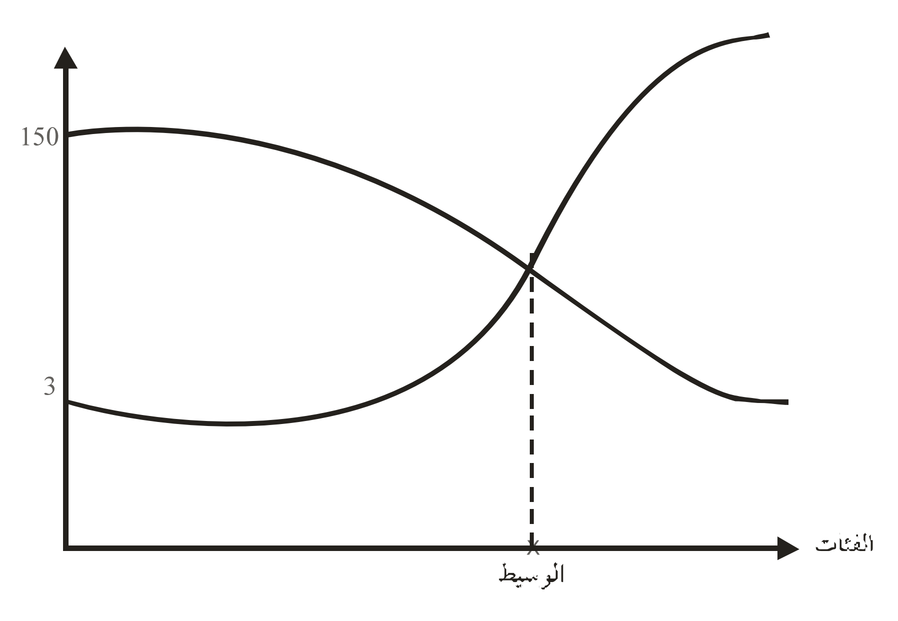
البرنامج التالي يقوم بحساب المتجمع التكراري الصاعد والمتجمع النـازل/الهـابط، وقـد صـمم البرنـامج، بطريقة عامة، يصلح لأي مجموعة مختلفة من البيانات، فقط تتغير قيمة N والبيانات في عبارة DATA.

برنامج لحساب المتجمع التكراري الصاعد والنازل.

```
10    REM
20    DIM  A(13), B(13), C(13), C(13), D(13),  E(13), F(13)
30    PRINT USING 250
40    PRINT USING 230
50    PRINT USING 240
60    PRING USING 250
70    PRINT
75    F1=0
80    READ N  REM NO OF OSERVATIONS
90    FOR I=1  TO N
100   READ A(I), B(I), F(I) REM الحد الأدنى، الحد الأعلى، والتكرار
110   F1 = F1 + F(I)
120   NEXT I
130   C(1) = F1
140   D(1) = F1
150   FOR I = 2 TON
160   C(I) = C(I-1) + F(I)
170   D (I) = D(I-1) – F (I-1)
180   NEXT I
190   FOR I = 1 TO N
200   E(I) = C (I) – D(I)
210   PRINT USING 260, E(I), D(I) , C(I), F(I), B(I), A(I)
220   NEXT I
222   PRINT
225   PRINT
230   الفئة   التكرار   المتجمع التكراري الصاعد   المتجمع التكراري النازل   الفرق
240   ________________________________________________
250   ****          ****          ****        ***        **.*-**.*
270   DATA          13, 12, 5,   15,5,2,    15.5 , 18.5,3,
280   18.5 ,   21.5 , 4 ,   21.5 ,    24.5 ,   5
290   24.5 ,   27.5,  14 ,   27.5 ,  30.5 , 21, 30.5
300   59,  33.5,  36.5  ,  24 ,  36.5 ,  39.5,  7,
310   39.5  ,  5,   42.5,   6,  42.5,  45.5,  2,
320   45.5  ,    48.5,    1,    48.5,    51.5 , 2
350   END
```

المخرجات:

الفئة	التكرار	المتجمع التكراري الصاعد	المتجمع التكراري النازل	الفرق
15.5-12.5	2	2	150	-148
18.5-15.5	3	5	148	-143
21.5-18.5	4	9	145	-136
24.5-21.5	5	14	141	-127
27.5-24.5	14	28	136	-108
30.5-27.5	21	49	122	-73
33.5-30.5	59	108	101	7
36.6-33.5	24	132	42	90
39.5-36.5	7	139	18	121
42.5-39.5	6	145	11	134
45.5-42.5	2	147	5	142
48.5-45.5	1	148	3	145
51.5-48.5	2	150	2	148



المتجمع الصاعد والمتجمع النازل

الوحدة الرابعة

مقاييس النزعة المركزية

MEASURES OF CENTRAL TENDENCY

النزعة المركزية: هو تجميع عدد كبير من المفردات حول قيمة متوسطة، ويقل عدد المفردات تدريجياً كلما بعدنا عن هذه القيمة المتوسطة، وتسمى هذه الظاهرة "بالنزعة المركزية".

ويوجد عدة مقاييس لقياس هذه النزعة أو القيمة المتوسطة مثل: المتوسط الحسابي، والوسيط...الخ.

1/4 المتوسط الحسابي **The Arithmetic Mean**

1/1/4 إيجاد المتوسط الحسابي للقيم المفردة "الفردية"

= مجموع القيم ÷ عدد القيم

فمثلاً: إذا كان لدينا القيم التالية 5, 7, 8, 10, 2, 4

فيكون المتوسط الحسابي لهذه القيم يساوي $\frac{5+7+8+10+2+4}{6} = \frac{36}{6} = 6$

فإن المتوسط الحسابي = $\frac{X_1 + X_2 + X_3 + ... + X_N}{N}$

X تشير إلى القيم.

و N تشير إلى عدد القيم.

قانون المتوسط الحسابي = $\overline{X} = \frac{\sum X}{N}$

حيث $\sum X = X_1 + X_2 + X_3 + ... + X_N$

إذا كانت قيم المشاهدات كبيرة، بحيث يصعب جمعها واستخراج المتوسط الحسابي بالطريقة السابقة، يمكن افتراض وسط فرضي كالآتي:

$$\overline{X} = A + \left(\frac{X_1 + X_2 + X_3 + ... + X_N}{N} \right)$$

ولنرمز للانحرافات عن الوسط الفرضي بالرمز d، وعلى ذلك فإن:

$$\overline{X} = A + \left(\frac{d_1 + d_2 + d_3 + ... + d_N}{N} \right)$$

$d_1 = (X_1+A)$

$d_2 = (X_2-A)$

$d_3 = (X_3-A)$

$d_N = (X_3-A)$

ويكون المتوسط الحسابي: $\overline{X} = A + \frac{\sum d}{N}$

حيث $\sum d = d_1 + d_2 + d_3 + ... + d_N$

مثـــال:

أوجد المتوسط الحسابي للقيم التالية مستخدماً طريقة الوسط الفرضي؟

5, 10, 15, 4, 6

الحل:

نفترض أن الوسط الفرضي يساوي +

$$\overline{X} = 6 = \frac{(5-6) + (10-6) + (15-6) + (4-6) + (4-6) + (6-6)}{5}$$

$$= \frac{(-1) + (4) + (9) + (-2)}{5} + 6$$

$$= \frac{10}{5} + 6$$

$= 8$

2/1/4 إيجاد المتوسط الحسابي لقيم مكررة

يكون حساب المتوسط الحسابي بضرب كل قيمة من القيم المكررة X في عدد تكراراتها، ثم قسمة ÷ حاصل الضرب على مجموع التكارات.

مثال:

احسب المتوسط الحسابي لأعمار موظفي شركة النصر للغزل والنسيج.

التكرار F	العمر X
3	20
4	27
6	32
7	36
20	المجموع

لكي يتم حساب المتوسط الحسابي، يجب ضرب القيم المكررة (العمر X في التكرارات (عدد التكرارات)، ثم قسمة حاصر الضرب ÷ على مجموع التكرارات...كالآتي:

Fx X	التكرار F	العمر X
60	3	20
108	4	27
192	6	32
252	7	36
612	20	المجموع

وبالتالي:

$$\overline{X} = \frac{\sum FX}{N} = \frac{612}{20} = 30.6$$

ويمكن استعمال طريقة الوسط الفرضي، وذلك بافتراض وسط فرضي، ثم إيجاد الانحرافات عنه، وبعد ذلك نقوم بضرب الانحرافات X في عدد تكراراتها، وجمع + حاصر الضرب، ثم قسمتها ÷ على مجموع التكرارات، كالآتي:

وبعد ذلك نأتي بالانحرافات عن الوسط الفرضي، كما مبين بالجدول التالي.

X	F	X-A=d	D × F
20	3	20 – 30 = -12	-36
27	4	27 – 32 = -5	-20
32	6	32 – 32 = 0	-
36	7	36 – 32 = 4	16
المجموع	20		-56 +28

$$\overline{X} = A + \frac{\sum dF}{N} \text{ أو } \overline{X} = A + \frac{\sum dF}{N} + A \quad -28$$

فالمتوسط الحسابي إذن يكون

حيث N هنا تشير إلى مجموع التكرارات، وبالتعويض.

$$\overline{X} = \frac{-28}{20} + 32$$

= -1.40 + 32

= 30.6

3/1/4 إيجاد المتوسط الحسابي لقيم مبوبة في فئات:

يحسب المتوسط الحسابي للتوزيعات التكرارية في فئات، حيث يتم استبدال الفئات المختلفة بمراكز الفئات، ويستعمل مركز الفئة فمثلاً لقيم مفردات الفئة وذلك لاعتباره تقديراً مناسبًا للمتوسط الحسابي الحقيقي لمفردات الفئة. ثم يجمع حاصل ضرب مركز الفئة في تكرارها نحصل على أقرب قيمة تقديرية لمجموعة قيم مفردات الفئات، وبقسمة هذا المجموع على مجموع التكرارات نحصل على أنسب تقدير للمتوسط الحسابي الحقيقي للقيم الأصلية.

إيجاد المتوسط الحسابي بالطريقة المطولة

الفئات	التكرار F	مركز الفئة X	FX
5-9	1	7	7
10-14	3	12	36
15-19	2	17	34
20-24	3	22	66
25-29	1	27	27
30-34	2	32	64
المجموع	12		234

$$\overline{X} = \frac{\text{مركز الفئة} \times \text{التكرار}}{\text{مجموع التكرارات}}$$

وبالتالي يمكن صياغة قانون المتوسط الحسسابي، كالآتي:

$$\overline{X} = \frac{\sum FX}{N}$$

$$\overline{X} = \frac{234}{12} = 19.5$$

ويمكن إيجاد المتوسط الحسابي باستخدام متوسط فرضي للمثال السابق:

نختار وسط فرضي من إحدى مراكز الفئات، ولـيكن مركـز الفئـة المقابـل الأكـبر تكـرار، ونجـد الانحراف عن الوسط الفرضي، ونطبق نفس الخطوات السابقة في إيجاد المتوسط الحسابي.

نفترض أن 22 وسط فرضي، والجدول التالي يبين ذلك.

الفئات	التكرار F	مركز الفئة X	X-A=d	D ×F
5-9	1	7	-15	-15
10-14	3	12	-10	-30
15-19	2	17	-5	-10
20-24	3	22	-	-
25-29	1	27	5	5
30-34	2	32	10	20
المجموع	12			-55 +25 -30

∴ إذن المتوسط الحسابي يكون $\overline{X} = \frac{\sum Fd}{N} + A$

$= \frac{-30}{12} + 22$

= 2.5 + 22

= 19.5

ويمكن إيجاد المتوسط الحسابي بالطريقة المختزلة، وذلك بأخذ عامل مشترك، وقسـمة الانحـراف على هذا العامل المشترك، وذلك كما مبين بالجدول التالي:

الفئات	F	X	d	$D=\frac{d}{k}$	D ×F
5-9	1	7	-15	-3	-3
10-14	3	12	-10	-2	-6
15-19	2	17	-5	-1	-2
20-24	3	22	-	-	-
25-29	1	27	5	1	1
30-34	2	32	10	2	4
المجموع	12				-11 5
					-6

إذن المتوسط الحسابي يكون: $\overline{X} = \frac{\sum F1d}{N} + xk + A$

$\overline{X} = \frac{-6}{12}.5 + 22$

$= \frac{-1}{2}.5 + 22$

$= \frac{-5}{2} + 22$

= -2.5 +22

= 19.5

4/1/4 خواص المتوسط الحسابــي:

1. مجموع انحرافات القيم عن وسطها الفرضي يساوي = صفر.

ويمكن إثبات ذلك جبرياً، فإذا كان

$$d_1 = X_1 - \overline{X}$$

$$d_2 = X_2 - \overline{X}$$

$$d_3 = X_3 - \overline{X}$$

$$d_N = X_N - \overline{X}$$

وبجمع الطرفين، ينتج أن:

$$\sum d = \sum X - N\,\overline{X}$$

ولكن

$$\overline{X} = \frac{\sum X}{N}$$

إذن صفر =

$$\sum = \sum X - N\left(\frac{\sum X}{N}\right)$$

نستنتج أنه إذا أخذنا انحرافات القيم أي رقم خلاف المتوسط الحسابي، فمجموع الانحرافات لا يساوي صفر.

2. إذا كان لدينا عدد من أزواج القيم المتغيرين Y,X لظاهرتين مستقلتين، فإن المتوسط الحسابي لمجموع قيم الظاهرتين معاً يساوي = مجموع المتوسط الحسابي لكل من الظاهرتين منفردتين، أي أنه، إذا كانت:

$$f_1 = X_1 \pm Y_1$$

$$f_2 = X_2 \pm Y_2$$

$$f_3 = X_3 \pm Y_3$$

$$f_N = X_N \pm Y_N$$

إذن $$\sum F = \sum X \pm \sum Y$$

وبالقسمة على N ينتج أن $$\overline{F} = \overline{X} \pm \overline{Y}$$

ويمكن تعميم هذه النتيجة في حالة أي عدد من المتغيرات.

3. مجموع مربعات انحرافات القيم عن وسطها الفرضي أصغر ما يمكن.
4. تدخل جميع القيم في حسابه.
5. لا يمكن حسابه في حالة التوزيعات التكرارية المفتوحة، وذلك لعدم قدرتنا على تحديد مركز الفئة المفتوحة.
6. لا يمكن حسابه في حالة البيانات النوعية.
7. يتأثر بالقيم المتطرفة.

5/1/4 تطبيقات مقاييس النزعة المركزية Measures of Central Tendency

1- الإحصائيـــة:

تعرف الإحصائية بأنها دالة العناصر (الوحدات) المكونة للعينة والتي لا تعتمد على معالم مجهولة، فالإحصائيات Statistics هي كميات يمكن حسابها من واقع قيم المفردات المكونة للعينة التي تم اختيارها، لذا فهي تكون الأساس الذي يرتكز عليه الاستنتاج الإحصائي أو الوصف الإحصائي للبيانات.

أنواع مقاييس النزعة المركزية

1. الوسط الحسابي $\frac{\sum X}{N} = M = \frac{S}{N}$ Arithmetic Mean=
2. الوسيط Median
3. المنوال Mode

4. الوسط الهندسي Geometric Mean
5. الوسط التوافقي Harmonic Mean

عدد القيم هنا يرمز له بالرمز = N

ومجموع المتغيرات بالرمز = S

والوسط الحسابي بالرمز = M

مثال: أوجد الوسط الحسابي للمفردات التالية:

23 , 13, 15, 19, 21, 21, 23, 20, 21, 22, 16

25 , 16, 22, 18, 29 , 20, 17

برنامج لحساب الوسط الحسابي لمجموعة مفردات

```
10    REM
20    S = 0
30    READ  N ,  REM                         عدد القيم
40    PRINT                                   البيانات
50    FOR I = 1  TO  N
60    READ X
70    PRINT, X
80    S = S + X
90    NEXT I
100   PRINT
110   PRINT
120   M = S/N
130   PRINT, M;  =                          الوسط الحسابي
140   PRINT
150   PRINT
160   DATA 17 , 16, 22, 21, 20, 23, 21, 19, 15, 13, 23, 17, 20, 29,
           18, 22, 16, 25
```

الوسط الحسابي، المخرجات

6/1/4 الوسط الحسابي المرجح بالتكرارات

أكثر الحالات التي يستخدم فيها الوسط الحسابي المرجح، هـي إيجـاد الوسـط الحسـابي للأسـعار المرتبطة بكميات مختلفة.

مثـال:

تقوم إحدى المؤسسات ببيع ثلاثة أنواع من السيارات، فإذا كان سعر السيارة من النوع الأول (أ) (ماركة مارسيدس) يساوي 25000 دينار، والسعر من الماركة (ب) (تويوتا) يساوي 23000 دينـار، وسـعر السيارة من الماركة (جـ) (نيسان Sunny) يساوي 18000 دينار.

أوجد متوسط سعر السيارة، علماً بأن الشركة قد باعت خلال العام المـاضي عـدد 1600 , 1400 , 1000 سيارة من الماركة (أ) و (ب) و (جـ).

الوسط الحسابي المرجح = $\frac{1600 \times 18000 + 1400 \times 2300 + 1000 \times 25000}{1600 + 1400 + 1000}$

$= \frac{86000000}{4000} = 21500$ دينار

فيما يلي برنامج لحساب الوسط الحسابي المرجح للبيانـات الـواردة في المثـال السـابق باسـتخدام المعادلتين: $A = \frac{T}{R}$, $B = \frac{S}{3}$

حيث:

الوسط الحسابي غير المرجح = A

$S = \sum_{i=1}^{3} Pi$

سعر السيارة P=

$T = \sum PiFi$

عدد السيارات F =

$R = \sum Fi$

برنامج لحساب الوسط الحسابي المرجح

```
10      REM
20      READ P1, F1, P2, F2, P3, F3
30      PRINT                 البيانات
40      PRINT
50      PRINT , P1
60      PRINT,  F1
70      PRINT , P2
80      PRINT,  F2
90      PRINT, P3
100     PRINT, F3
110     S = P1 + P2 + P3
120     T = P1 * F1 + P2*F2 + P3*F3
130     R = F1 + F2 + F3
140     A = S/3
150     B = T/R
160     PRINT
170     PRINT, a;` =                  الوسط الحسابي غير المرجح
190     PRINT
200     PRINT, B; ' =                 الوسط الحسابي المرجح
210     PRINT
220     PRINT
230     DATA 25000, 1000, 23000, 1400, 18000, 1600
240     END
```

المخرجات: **البيانات**

25000
1000
23000
1400
18000
1600

الوسط الحسابي غير المرجح = 22000

الوسط الحسابي ا لمرجح = 21500

2/4 الوسيـــط The Median

الوسيط هو القيمة التي يسبقها، ويليها مجموعتين متساويتين من المشاهدات. فالوسيط: كلمة تعني منتصف الشيء، وهو القيمة التي تقع في منتصف القيم المعطاه، وذلك بعد أن يتم ترتيبها إما تصاعدياً أو تنازلياً.

1/2/4 إيجاد الوسيط في حالة قيم غير مبوبة

حساب الوسيط في حالة القيم غير المبوبة، بترتيب المشاهدات الموجودة إما تصاعدياً أو تنازلياً.

يمكن إيجاد الوسيط كما يلي = القمية التي ترتيبها $\frac{N+1}{2}$

حيث N تشير إلى عدد القيم.

- إذا كان عدد القيم فردياً، فإن قيمة الوسيط تكون هي قيمة المشاهدة الواردة في هذا المكان.
- إذا كان عدد القمي زوجياً، فإن القيمة الوسيط تكون قيمة متوسط بين القيمتين المركزيتين.

مثلاً إذا كان لدينا عدد القيم فرديا، وكانت لدينا القيم التالية:

10 , 16, 20, 14, 30 , 26 , 22

- إن عدد القيم فردي هو عبارة عن 7 مفردات.
- نرتب القيم تصاعدياً، كالآتي:
- ثم نطبق القانون السالف الذكر $\frac{7+1}{2} = \frac{N+1}{2} = 4$

30 , 26, 22, <u>20</u>, 16, 14, 10

- الوسيط هو القيمة الرابعة من القيم = 20
- في حالة ما إذا كانت القيم زوجية، مثل 52 ,44 ,40 ,32 ,24 ,14 ,12 ,8.

$$\frac{N+1}{2} = \frac{8+1}{2} = 4.5$$

∴ قيمة الوسيط تقع بين القيمتين الرابعة والخامسة وهي القيم 24 , 32.

∴ إذن الوسيط = $28 = \frac{56}{2} = \frac{32+24}{2} =$

2/2/4 حساب الوسيط في حالة البيانات المبوبة في فئات

يمكن حساب الوسيط في حالة البيانات المبوبة في فئات بأن:

الخطوة الأولى:

- نجد مجموع التكرارات في التوزيع.
- ثم نجد ترتيب الوسيط وذلك بقسمة ÷ مجموع التكرارات ÷ على 2.

الخطوة الثانية:

- هو إيجاد التكرارات المتجمع الصاعد أو النازل.
- ثم نحدد الفئة التي يقع ضمنها الوسيط.

الوسيط عبارة عن الحد الأدنى للفئة التي يقع ضمنها الوسيط، ونضيفه للمعلومات التي تحصلنا عليها.

أما إذا استخدمنا التكرار المتجمع النازل فيمكن طرح البيانات التي تحصلنا عليها من الحد الأعلى للفئة التي يقع ضمنها الوسيط.

استخدم متغيراً منفصلاً في المثال التالي:

الفئات	F	التكرار المتجمع الصاعد
5-9	3	3
10-14	2	5
15-19	6	11
20-24	10	21
25-29	13	34
30-34	7	41
35-39	4	45
المجموع	45	

التكرارات التي تسبق الفئة الوسيطية ← 21

القيمة التي تحتوي الوسيط ← 41

$$\text{ترتيب الوسيط} = \frac{N}{2}$$

$$\frac{45}{2} = 22.5$$

الفئة التي تتضمن الوسيط هي 25-29

حدها الأدنى 25 وحدها الأعلى 29 ومداها = 5

والحد الأدنى الحقيقي 24.5، والحد الأعلى الحقيقي 29.5 وعدد تكراراتها = 13

ويمكن التعبير عن الوسيط بالقانون التالي:

حيث M يعني الوسيط: $M=L+\left(\frac{P}{F}\right)\mathbf{R}$

و L تشير إلى الحد الأدنى الحقيقي للفئة الوسيطية.

و F تشير إلى تكرارات الفئة التي تحوي الوسيط.

و I تعني مدى الفئة.

و P تساوي $\frac{N}{2} - C$

حيث N تشير إلى مجموع التكرارات في التوزيع($\sum F$)

و C تشير إلى عدد التكرارات في المتجمع الصاعد التي تسبق الفئة الوسيطية.

ويمكن تطبيق هذا القانون على البيانات الموجودة في المثال لإيجاد الوسيط، وذلك على النحو التالي:

$$M = 24.5 + \left[5.\frac{\frac{45}{2} - 21}{13} \right]$$

$$= 24.5 + \left(5.\frac{1.5}{13} \right)$$

$$= 24.5 + .58$$

$$= 25.08$$

ويمكن إيجاد الوسيط باستعمال التكرار المتجمع النازل، كالآتي:

الفئات	F	التكرار المتجمع النازل
5-9	3	45
10-14	2	42
15-19	6	40
20-24	10	34
25-29	13	24
30-34	7	11
35-39	4	4
المجموع	45	

ويمكن استعمال نفس الخطوات السابقة، إلا أن قانون الوسيط سيكون:

$$M = L\left(\frac{P}{F}\right)^{1}$$

حيث L هنا تعني الحد الأعلى الحقيقي للفئة الوسيطية.

وحيث أن الوسيط يقع في الفئة 25 – 29 ومداها 5 وعدد تكراراتها 13، وبالتالي يكون الوسيط:

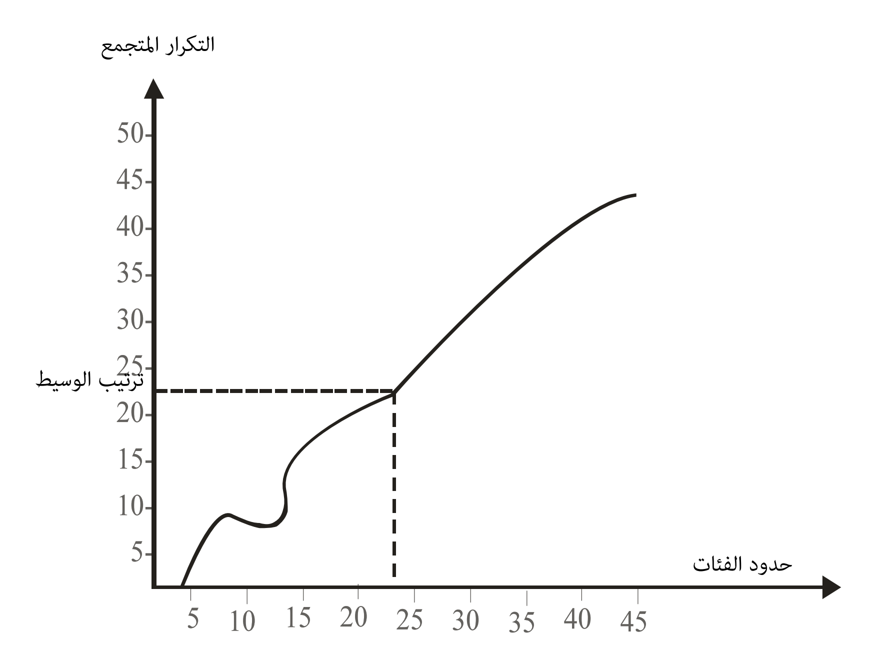
$$M = 29.5 = \left(\frac{\frac{45}{2}}{13} - 11\right)$$

$$= 29.5 - 4.42$$

$$= 25.08$$

كما يمكن إيجاد الوسيط بالرسم البياني، وبالتطبيق على المثال السابق وباستعمال المتجمع الصاعد.

إيجاد الوسيط بواسطة الرسم البياني.



كما يمكن إيجاد الوسيط برسم المنحنى المتجمع الصاعد والنازل في رسم واحد، وعند تقاطع المنحنى نسقط عمود على المحور الأفقي، وتكون هذه القيمة هي الوسيط.

3/2/4 خواص الوسيط

1. يلاحظ أن حساب الوسيط يعتمد على التكرارات لا على مراكز الفئات، لذا يمكن استخدامه في الفئات المفتوحة.
2. لا يتأثر بوجود قيم متطرفة.
3. يمكن حسابه للبيانات النوعية.
4. مجموع الانحرافات المطلقة للقيم عن وسيطها أقل ما يمكن.

4/2/4 تطبيقات الوسيط على الحاسوب

الوسيط Median

الوسيط هو القيمة الوسطى للمقادير المرتبة ترتيباً تصاعدياً أو تنازليًا كما أن الوسط الحسابي للقيمتين الوسيطتين إذا كان عدد المتغيرات زوجياً.

البرنامج التالي يقوم بحساب الوسيط للمفردات، والبرنامج يحتوي عـلى فقـة لفـرز القـيم تنازليـاً (السطر 45 إلى السطر 110).

برنامج لحساب الوسيط لمجموعة مفردات:

```
10    REM
20    DIM  X  (17)
25    READ N REM   عدد القيم
30    FOR  I = 1 TO N
35    READ X (I)
40    NEXT I
45    REM  فرز القيم تنازلياً
50    FOR I = 1 TO N-1
60    FOR J = I + 1  TO N
65    IF  X (I)  > X (J)  THEN 100
70    T = X(I)
80    X(I) = X(J)
90    X(J) = T
100   NEXT J
110   NEXT I
120   PRINT,  القيمة، الترتيب
130   PRINT,  القيمة، الترتيب
140   FOR I = 1 TO N
150   PRINT, I , X(I)
160   NEXT I
170   PRINT
180   P = INT ((N+1) /2)
190   M = X (P)
      200    PRINT M                الوسيط هو المتغير رقم ., P، وهو الرقم., M
210   DATA  17,25,16,36,17,18,19,16,22,19,38,41,21,36,81,17,44,13
```

المخرجات

الترتيب	القيمة	
1	81	
2	44	
3	41	
4	38	
5	36	
6	36	
7	25	
8	22	
9	21	← Median
10	19	
11	19	
12	18	
13	17	
14	17	
15	16	
16	16	
17	13	

Median الوسيط هو المتغير رقم 9 وهو الرقم 21.

4/3 الربيع الأعلى والربيع الأدنى: Quartiles

تبين مما سبق أن الوسيط، بأنه يقع في منتصف القيم، إذا كانت القيمة من التوزيع يسبقها 25% من المشاهدات في التوزيع، ففي هذه الحالة يطلق على المقياس الذي ينتج بالربيع الأدنى، ويمكن تطبيق قانون الوسيط في إيجاد الربيع الأدنى، إلا أن ترتيبه يتم بقسمة مجموع التكرارات ÷ على 4 أو N/4، والمثال يوضح كيفية إيجاده.

الفئات	التكرار F	التكرار المتجمع الصاعد
20 – 24	3	3
25 – 29	9	12
30 – 34	13	25
35 – 39	16	41
40 – 44	20	61
45 – 49	15	77
50 – 54	12	89
55 – 59	8	97
60 – 64	4	100
المجموع	100	

يمكن تطبيق نفس قانون الوسيط، مع الفارق في ترتيب الربيع الأدنى، فإذا ما رمزنا للربيـع الأدنى بالرمز Q_1، أما باقي الرموز فتشير إلى نفس المعنى السابق المستخدم في قانون الوسـيط، وبالتـالي يمكـن التعبير عن الربيع الأدنى، بالقانون:

$$Q1 = L + \left(\frac{P}{F}\right)^{1}$$

حيث
$$P = \frac{N}{4} - C$$

ويمكن تطبيق هذا القانون على المثال السابق، فيكون ترتيب الربيع الأدنى $\frac{100}{4}$ أو 25، وبالتـالي فإن الفئة التي يقع ضمنها الربيع الأدنى هي الفئة 34-30.

ومدى هذه الفئة 5 وبالتالي، فإن الربيع الأدنى، يكون:

$$Q1 = 29.5 + \left(\frac{\frac{100}{4} - 12}{13}\right)^{.5}$$

$$= 29.5 + \left(\frac{25 - 12}{13}\right)^{.5}$$

$$= 29.5 + \frac{13.5}{13}$$

$$= \boxed{34.5}$$

ويمكن إيجاد الربيع الأدنى عن طريق الرسم كما في الوسيط.

أما إذا كانت القيمة في التوزيع يسـبقها 75% مـن عـدد المشـاهدات في التوزيـع التكـراري، فـإن المقياس هنا يطلق عليه الربيع الأعلى، ويمكن إيجاده بنفس الطريقة السابقة إلا أن ترتيبه يتم إيجـاده بضرب مجموع التكرارات في 3 وقسمتها ÷ على 4.

ويمكن استخدام نفس المثال السابق الذي استخدم في إيجاد الربيع الأدنى، فترتيب الربيع الأعـلى هو $\frac{3 \times 100}{4}$ أي 75 وهذا يقع في الفئة 49-45 ومداها 5 وتكراراتها 15، ويمكن تطبيـق نفـس القـانون السابق.

وإذا رمزنا للربيع الأعلى بالرمز Q_2، فإنه يكون:

$$Q_2 = L + \left(\frac{P}{F}\right)^1$$

$$P = \frac{3N}{4} - C$$

$$Q_2 = 44.5 + \left(\frac{\frac{3.100 - 16}{4}}{15}\right)^5$$

$$Q_2 = 44.5 + \left[\frac{\frac{3.100}{4} - 16}{15}\right]^5$$

$$P = \frac{3N}{4} - C$$

حيث:

إذن

$$Q_2 = 44.5 + \left(\frac{\frac{3.100 - 16}{4}}{15}\right)^5$$

$$= 44.5 + \frac{14.5}{15}$$

$$= \boxed{48.5}$$

كذلك يمكن إيجاد الربيع الأعلى، عن طريق الرسم، وذلك بنفس الطريقة التي يتم بها إيجاد الوسيط عن طريق الرسم.

3/4 تطبيقات الربيع الأعلى، والأوسط، والأدنى على الحاسوب

1/3/4 الربيعات (الربيع الأعلى، والأوسط، والأدنى) Quartiles

الربيعات هي التي تقسم القيم إلى أربعة أقسام يساوي كل منها الربع (25%).

أ- الربيع الأعلى (و 75%).

وهو الذي يقسم القيم إلى جزأين بحيث يكون عدد القيم التي أقل منه يساوي = $\frac{3}{4}$، والربع الباقي أكثر منه.

ويمكن استخراجه من البيانات المبوبة حسب القاعدة:

$$(,75\%) = ح_1 + \frac{\left(\frac{3}{4}ن - \bar{ن}\right) ط}{ك}$$

حيث:

ح1: هي فئة الربيع الأعلى التي يعلو عندها التجمع التكراري الصاعد لقيمة $\frac{3}{4}$ن لأول مرة.

$\bar{ن}$: هو المتجمع التكراري الصاعد لدى الفئة التي تسبق فئة الربيع الأعلى.

ط: هي طول فئة الربيع الأعلى.

ك: هي تكرار فئة الربيع الأعلى.

ويلاحظ أن القاعدة نفسها هي قاعدة الوسيط مع اختلاف تفسير الرموز.

ب- الربيع الأوسط (, 50%) ، وهو الوسيط Median.

جـ- الربيع الأدنى (, 25%).

وهو القيمة التي تعلو $\frac{1}{4}$ القيم بينما تعلو عليها $\frac{3}{4}$ تلك القيم، وبذلك تكون قاعدة الربيع الأدنى هي: ,25% =

$$,25\,\% = ح_1 + \frac{\left(\frac{ن}{4} - \bar{ن}\right) ط}{ك}$$

حيث:

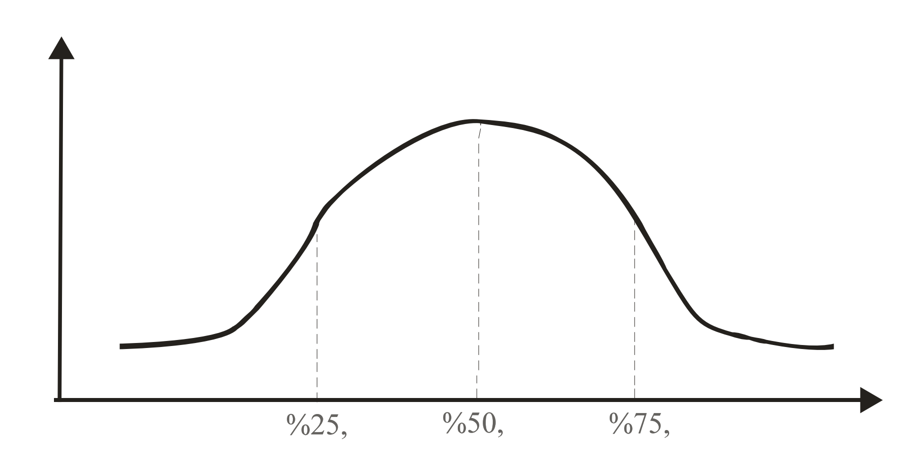
ح1= هو الحد الأدنى لفئة الربيع الأدنى التي يعلو عندها التجمع التكراري الصاعد لقيمة $\frac{ن}{4}$ لأول مرة.

ن̄ = هو المتجمع التكراري الصاعد لدى الفئة التي تسبق فئة الربيع الأدنى.

ط= هو طول فئة الربيع الأدنى.

ك= تكرار فئة الربيع الأدنى.

% 25, < %50, < % 75 ,



مثال: استخدم البيانات الواردة أدنا لاستخراج الربيع الأعلى والربيع الأدنى:

رقم الفئة	الفئات العمرية	كِ تكرار F	المتجمع الصاعد	
1	12.5 – 15.5	2	2	
2	15.5 – 18.5	3	5	
3	18.5 – 21.58	4	9	
4	21.5 – 24.5	5	14	
5	24.5 – 27.5	14	28	
6	27.5 – 30.5	21	49	فئة الأدنى
7	30.5 – 33.5	59	108	فئة الوسيط
8	33.5 – 36.5	24 F	132	فئة الأعلى
9	36.5 – 39.5	7	139	
10	39.5 – 42.5	6	145	
11	42.5 – 45.5	2	147	
12	45.5 – 48.5	1	148	
13	48.5 – 51.5	2	150	
	$\sum$	150		المجموع

$112.5 = \frac{150\times 3}{4} = \frac{3}{4}\text{ن} : \%\ 75,$

$$\frac{3\times\left(108 - 112.5\right)}{24} + 33.5 = \%\ 75, \quad \therefore$$

$34\ .063 = \%\ 75,\ \therefore$

$37.5 = \frac{\text{ن}}{4} : \%25,$

$$\frac{3\times\left(28 - 37.5\right)}{21} + 27.5 = \%\ 25, \quad \therefore$$

$28\ .857 = \%\ 52,\ \therefore$

البر نامج التالي يقوم باستخراج الآتي، مستخدماً البيانات السابقة:

- فئة الربيع الأعلى.
- الفئة الوسيطية.
- فئة الربيع الأدنى.
- الربيع الأعلى.
- الوسيط.
- الربيع الأدنى.

باستخدام المعادلة العامة:

$$Q = L(1) + \frac{\left(S(1)\ -\ 0(1)\right)P(1)}{W(1)}$$

وتكون القمية هي الربيع الأعلى أو الوسيط أو الربيع الأدنى، عندما تكون 1 تسـاوي 1 أو 2 أو 3 على التوالي.

حيث:

Q= الربيع

L= الحد الأدنى للفئة

S= النسبة من عدد المتغيرات

D= المتجمع التكراري للفئة السابقة

P= طول الفئة

W= تكرار الفئة

برنامج لحساب الربيعات لبيانات متجمعة

```
10      REM
20      DIM A(13), B(13) , C(13), C(13), D(13), E(13), F(13), G(13)
30      T=0
40      READ N
50      FOR I = 1 TO N
60      READ A (I), B(I), F(I)
70      T = T+F(I)
80      NEXT I
90      C(1) =  F(1)
100     FORI = 2 TO N
110     C (I) = C (I-1) + F(I)
120     NEXT I,
125     PRINT, ________________________________________
130     PRINT,       رقم الفئة   الفئة  التكرار   المتجمع التكراري الصاعد
140     PRINT, ________________________________________
150     PRINT
160     PRINT
170     FOR I = 1 TO N
180     PRINT USING 190, C(I) , F(I), B(I), A(I), I
190     ***     **     **.**  -  **.** **
200     NEXT I,
210     PRINT
220     PRINT
230     PRINT TAB (33); T; TAB (50); المجموع
240     PRINT
250     PRINT
260     S1 = TX 3/4
270     S2 = T /2
280     S3 = T/4
290 FOR I = 1 TO N
300     IF C(I) > S1 THEN 320
310     J = I
320     NEXT I
330     FOR I = 1 TO N
340     IF C(1) > S2 THEN 360
350     K = i
360     NEXT 1
370     FOR I = 1 TO N
380     IF C(I) > S3  THEN 400
390     V = I
400     NEXT I
```

```
410    PRINT USING 470, B (J+1), A (J+1), J+1
420    PRINT
430    PRINT USING 480, B (K+1) , A (K+1) , K+1
440    PRINT
450    PRINT USING 490, B (V+1), A(V+1), V+1
460    PRINT
470    فئة الربيع الأعلى هي الفئة رقم   **وهي  *.** - *.**
480    فئة الوسيطية هي الفئة رقم        **وهي  *.** - *.**
490    فئة الربيع الأدنى هي الفئة رقم   **وهي  *.** - *.**
500    GO SUB  630
510    Q1 = FNA (L1, S1, D1, P1, W1)
520    PRINT, Q1; = الربيع الأعلى
  5    PRINT
530    GO SUB 680
540    Q2 = FNA (L2, S2, O2, P2, W2)
550    PRINT, Q2;  = الوسيط
560    PRINT
570    GO SUB 730
580    Q3 = FNA (L3, S3, O3, P3, W3)
590    PRINT, Q3; =  الربيع الأدنى
600    PRINT
610    PRINT
620    STOP
630    L1 = A (J+1)
640    O1 = C(J)
650    P1 = B (J) – A(J)
660    W1 = F(J+1)
670    RETURN
680    L2 – A (K+1)
690    O2 = C (K)
700    P2 = B (K) – A(K)
710    W2 = F (K+1)
720    RETURN
730    L3 – A (V+1)
740    O3= C(V)
750    P3 = B(V) – A(V)
760    W3 = F (V+1)
770    RETURN
750    DEF FNA (L,X,O,P,W)=
```

```
        = L + ( (X-0) * P) /W
810     DATA
900     END.
810     DATA 13, 12.5, 15.5, 2, 15.5
820     18.5, 3, 18.5, 21.5, 4, 21.5,
830     24.5, 5, 24.5, 27.5, 14, 27.5,
840     30.5, 21, 30.5, 33.5, 59, 33.5,
850     36.5, 24, 36.5, 39.5, 7, 39.5, 42.5
860     6, 42.5, 45.5, 2, 45.5, 48.5, 1,
870     48.5, 51.5, 2
900     END
```

المخرجات

رقم الفئة	الفئة	التكرار F	المتجمع التكراري الصاعد
1	15.5 – 12.5	2	2
2	18.5 – 15.5	3	5
3	21.5 - 18.5	4	9
4	24.5 – 21.5	5	14
5	27.5 – 24.5	14	28
6	30.5 – 27.5	21	49
7	33.5 – 30.5	59	108
8	36.5 – 33.5	24	132
9	39.5 – 39.5	7	139
10	42.5 – 39.5	6	145
11	45.5 – 42.5	2	147
12	48.5 – 45.5	1	148
13	21.5 – 48.5	2	150
	المجموع	150	

- فئة الربيع الأعلى هي الفئة رقم 8 وهي 33.5 -36.5
- الفئة الوسيطية هي الفئة رقم 7 وهي 30.5 – 33.5
- فئة الربيع الأدنى هي الفئة رقم 6 وهي 27.5-305

الربيع الأعلى = 34.0625

الوسيط = 31.82202

الربيع الأدنى = 28.85713

4/4 المنـــــوال: The Mode

المنوال هو القيمة التي تتكرر أكثر من غيرهـا، مـثلاً إذا كانـت لـدينا قـيم للمشـاهدات التاليـة: 6,8,14,8,6,12,8,10

- فإن المنوال لهذه المشاهدات يكون مساوياً للقيمـة 8 وذلـك لأنهـا أكـثر تكـراراً بـين المشـاهدات السابقة.
- أما إذا وجدت قيمة مكررة نفس العدد، فإنه يمكن أن يكون هناك منوالان لهذه المشاهدات.
- أما في الجداول المبوبة في فئات، فإن القيمة المنوالية هي الفئة ذات التكرار الأكبر مـن أيـة فئـة أخرى في التوزيع.

مثـــال:

	مركز الفئة X	التكرار F	الفئات
	7	2	5 – 9
	12	5	10 – 14
	17	10	15 – 19
الفئة المنوالية	22	17	20 – 24
	27	15	25 – 29
	32	8	30 – 39
	37	3	35 - 9

أكبر تكرار في هذه الحالة هو تكرار الفئة 24-20 وهذه هي الفئة المنوالية.

فالمنوال = يساوي مركز الفئة المقابل لهذه الفئة، أي أن المنوال = 22.

هذه الطريقة تقريبية وبدائية جداً. وذلك لأن الموال موزع عـلى الفئـة كلهـا (مـن القيمـة 19.5 حتى القيمة 24.5). وبالتالي يجب أن يقدر المنوال بصورة أدق.

وهناك طريقتان لتحقيق ذلك هما:

(أ) طريقة الرافعة.

(ب) طريقة الفروق (بيرسون Person)

1/4/4 طريقة الرافعـــة

تأخذ هذه الطريقة في الاعتبار، الفئة المنوالية، وتكراري الفئتين المحيطين بها، على أساس أن تكراري هاتين الفئتين يتجاذبان المنوال، فتميل قيمة المنوال للاقتراب من الفئة ذات التكرار الأكبر.

مثـال:

الفئات	التكرار F
15 – 19	10
20 – 24	17
25 - 29	15

19.5 الحد الأدنى للفئة المنوالية × (طول الفئة – X)

24.5 الحد الأعلى للفئة المنوالية

تكرار الفئة التي تسبق الفئة المنوالية (10)

المنوال

تكرار الفئة التي تلحق الفئة المنوالية (15)

والفئة التي تحتوي المنوال هي الفئة 24-20 وطول الفئة يساوي 5 وحدّها الأدنى الحقيقي، يساوي 19.5 وحدّها الأعلى الحقيقي يساوي 24.5.

ولكي يتم حساب المنوال، نفرض أن هناك رافعة طولها يساوي مدى الفئة (5):

أي طول الفئة المنوالية ونقطة ارتكازها يقع على بعد X من الحد الأدنى للفئة المنوالية، ويقع على بعد طول الفئة -5 أي (5-X) من الحد الأعلى للفئة المنوالية.

وبتطبيق قانون الروافع:

القوة × ذراعها = المقاومة × ذراعها

إذن (5-X) 15 = 10 X "فك الأقواس"

3 = X "بالقسمة على ÷ 25"

إن المنوال يساوي الحد الأدنى الحقيقي للفئة المنوالية + قيمة × المنوال

المنوال يكون: M=L+X   M = 19.5 + 3

= 22.5

2/4/4 طريقة الفروق (بيرسون): **Pearson**

تستعمل هذه الطريقة الفروق بين التكرارات المحيطة بالفئـة المنواليـة وتكـرار الفئـة المنواليـة، وتقسم ÷ هذه الطريقة مدى الفئة التي يقع ضمنها المنوال إلى ÷ قسمين بنسبة هذين الفرقين.

ولو استخدمنا المثال السابق، فإننا نجد أن الفرق بين تكرار الفئة المنوالية وتكرار الفئـة السـابقة لها يساوي =7 والفرق بين تكرار الفئة المنوالية وتكرار الفئة اللاحقة لها يساوي 2. ولو فرضنا أن القسم الأول المناظر للمنوال طواله يساوي X فإن طول القسم الثاني يساوي = (5-X)، وعلى ذلك فـإن النسـبة بين هذين القسمين تكون:

الفئات	التكرار F	الفرق
15 – 19	10	7
20 – 24	17	17
25 – 29	15	2

وبضرب الطرفين X في الوسطين ينتج:

$$\frac{X}{5-X} = \frac{7}{2}$$

$$2X = 35 - 7X$$

$$9X = 35$$

$$\therefore \quad X = \frac{35}{9} = 3.9$$

ومن هنا يكون المنوال، كالآتي:

$$D = 19.5 + 3.9$$

$$= 23.4$$

ويلاحظ أن قيمة المنوال مختلفة، وهذا متوقع، نظراً لأن الطرق المتبعة في إيجاد المنوال مختلفة، إلا أن طريقة الفروق هي أدق الطرق في إيجاد المنوال، وذلك لأنه إذا كان الفرق صغيراً بين تكرار الفئة المنوالية وتكرار فئة مجاورة لها فهنا يعني أن تكرار هذه الفئة أقرب إلى أكبر تكرار، وهو تكرار المنوال، وعلى ذلك، فالمنوال يكون أقرب إلى هذه الفئة المجاورة مما لو كان الفرق كبيراً.

3/4/4 تطبيقات المنوال على الحاسوب Mode

مثال على المنوال

أوجد المنوال في كل من الحالات التالية:

(أ) 19, 26, 24, 10, 12, 19, 19, 18, 16, 16, 14

(ب) 26, 24 10, 12, 19, 19, 18, 18, 16, 16, 14

(ج) 26, 24, 10, 12, 19, 18, 16, 14

الحل:

(أ) المنوال = 19 لأنها الأكثر تكراراً.

(ب) لها ثلاثة منوالات وهي = 16 , 18 , 19

(ج) ليس لها منوال.

فيما يلي برنامج لحساب المنوال لبيانات تكرارية وباستخدام المعادلة:

طريقة الفروق Pearson

حيث

$$M = L + \left(\frac{F1}{F1+F2}\right).C$$

M = المنوال

L = الحد الأدنى للفئة المنوالية

F1= الفرق بين تكرار الفئة المنوالية والسابقة عليها

F2= الفرق بين تكرار الفئة المنوالية واللاحقة لها

برنامج لحساب المنوال لبيانات متجمعة

```
10      REM
20      DIM A(13) , B(13) , F(13)
30      T = 0    REM                 مجموع التكرارات
40      READ   N  REM                  عدد القيم
50      PRINT USING 70
60      PRINT USING 80
70                                  التكرار       الفئة
80                                  ______      _____
85      PRINT
90      FOR I = 1 TO N
100     READ A (I), B (I), F(I) REM الحد الأدنى، الحد الأعلى، التكرار
110     T = T + F  (I)
115     PRINT USING 130, F(I), B(I), A(I)
120     NEXT I
130     :   **      **.*   -   **.*
140     PRINT
150     PRINT USING 170
160     PRINT USING 180, T
170     :           ________
180     :           ***                                  المجموع
190     REM                                        لإيجاد التكرار الأعلى
200     H = F (1)
210     FOR I = 2  TO  N
220     IF  F (I) < H  THEN 250
230     H = F (I)
240     J = I
250     NEXT I
260     F1 = F (J) – F(J-1)
270     F2= F (J) – F (J+1)
280     L = A (J)
```

```
290     C = B (J) – A (J)
300     M = L + F1/ (F1+F2) *C
310     PRINT
320     PRINT USING 325, B(J) , A(J)
325     :    **.*  - **.*
330     PRINT
340     PRINT , M; ` =                                     المنوال
350     PRINT
360     DATA    13,12,5,15.5, 2, 15.5, 18.5, 3,
370     DATA   18.5, 21.5, 4, 21.5, 24.5, 5, 24.5, 27.5,
380     DATA   14, 27.5 , 30.5, 21.5, 30.5, 33.5, 59, 33.5
390     DATA    36.5, 24, 36.5, 39.5, 7, 39.5, 42.5
400     DATA    6, 42.5, 45.5, 2, 45.5, 48.5, 1, 48.5
        DATA 51.5, 2.
410     END
```

المخرجات

الفئة	التكرار F	
12.5 – 15.5	2	
15.5 – 18.5	3	
18.5 – 21.5	4	
21.5 – 24.5	5	
24.5 – 27.5	14	
27.5 – 30.5	21	
30.5 – 33.5	59	الأكثر تكراراً هو المنوال
33.5 – 36.5	24	
36.5 – 39.5	7	
37.5 – 42.5	6	
42.5 – 45.5	2	
45.5 – 48.5	1	
48.5 – 51.5	2	
	150	المجموع

الفئة المنوالية 33.5 – 30.5

المنوال =   32.06163

4/4/4 خصائص المنوال واستخداماته

1. يمكن استخدامه لاستخراج مقياس أفضل للنزعة المركزية إذا كانت البيانات وصفية.
2. يستخدم كمقياس تقريبي للنزعة المركزية، لعدم قابليته العمليات الجبرية، إلا أنه يسهل تقريبه، سواء بطريقة بيرسون أو بمركز الفئة المنوالية.
3. لا يتأثر بالقيم المتطرفة، ويمكن استخراجه في حالة الفئات المفتوحة التي يندر أن تكون فئات منوالية.

العلامة بين الوسط والوسيط والمنوال:

$$\text{الوسط الحسابي - الوسيط} = \frac{(\text{الوسط الحسابي - المنوال})}{\sum}$$

4. أسهل مقاييس النزعة المركزية.
5. لا يتأثر بوجود قيم متطرفة.
6. يمكن حسابه في حالة التوزيعات التكرارية المفتوحة، بحيث لا تكون الفئة المنوالية إحدى الفئات المفتوحة.
7. يمكن حسابه للبيانات النوعية.
8. يكون لبعض البيانات منوالات وأكثر.

5/4 الوسط الهندسي **The Geometric Mean**

يستخدم الوسط الهندسي في حالة المتغير الذي يكون في شكل نسب أو أرقام قياسية.

استعمال الوسط الهندسي أقل من المقاييس الأخرى نظراً لاستخدام اللوغاريتمات في حساب هذا الوسط الهندسي مما يجعل حسابه معقداً، ومن جهة ثانية نظراً لاستعمال اللوغاريتمات في حسابه، فإنه من الصعب شرحه.

مثال: أوجد الوسط الهندسي للآتي:

وحدات السلعة	التكلفة	لو
1	1.50	.1761
2	2.10	.3222
3	1.258	.5969
4	1.75	.2430
5	1.90	32788
		1.1170

والفرق بين حساب الوسط الهندسي والوسط الحسابي هو أن القيم تحول إلى لوغاريتمات في حالة الوسط الهندسي، ومن ثم إيجاد مقابل اللوغاريتم.

وفي هذا المثال يمكن إيجاد اللوغاريتم (لو) لكل وحدة من السلعة، ثم جمع اللوغاريتمات وتقسيمها ÷ على عددها (5)، ثم إيجاد العدد المقابل للوغاريتم.

في المثال السابق مجموع لو 1.1170 ، ثم تقسيم هذه اللوغاريتمات على (5) أي تحصل على متوسط:

$$\frac{1.1170}{5} = .2234$$

ثم نجد مقابل العدد وهو 1.67

وهذا هو الوسط الهندسي 1.67، ويلاحظ أن ا لمتوسط الحسابي لهذه القيم في هذا المثال يساوي 1.70... وهذا يعني أن المتوسط الهندسي أقل من الوسط الحسابي، وذلك لأن القيم الكبرى ليس لها تأثير على الوسط الهندسي.

ويمكن القول بصفة عامة أن قانون المتوسط الهندسي في حالة القيم غير المبوبة تكون كالآتي:

$$GM = \frac{\sum \text{Log} \ x}{N}$$

أما البيانات المبوبة في توزيع تكراري، فإن قانون المتوسط الهندسي يكون:

$$GM = \frac{\sum F \ \text{Log} \ x}{N}$$

وفي الحالتين يجب إيجاد مقابل اللوغاريتم.

1/5/4 تطبيقات الوسط الهندسي على الحاسوب

الوسط الهندسي The Geometric Mean

هو الجذر النوني لحاصل ضرب N ن قيمة عينية، أي أن الوسط الهندسي للمتغيرات Xn , ... , X_3, X_2, X_1 هو

$$GM = N\sqrt{X_1 \times X_2 \times X_3 \ .... * Xn}$$

$$= \frac{\text{Log } x_1 + \text{Log } x_2 + \text{Log}x_3 + ... + \text{Logxn}}{N}$$

إذاً لوغاريتم الوسط الهندسي للقيم هو الوسط الحسابي للوغاريتمات تلك القيم.

مثال: أوجد الوسط الهندسي للقيم.

23, 13, 15, 19, 21, 23, 20, 21, 22, 16

25, 16, 22, 18, 29, 20, 17

$$GM = \frac{\sum \text{Log } x}{N} = 19.625$$

البرنامج أدناه يقوم بحساب الوسط الهندسي لمفردات، باستخدام المعادلة.

حيث:

$G = \sqrt[N]{T}$

G= الوسط الهندسي

N = عدد المتغيرات

$T = T1 \times T2 \times T3 \times ... \times TN$

برنامج لحساب الوسط الهندسي لمجموعة مفردات

```
10      REM
20      T = 1
25      PRINT  البيانات
27      PRINT  , ____________,
30      READ N REM  عدد الأرقام
40      FOR I = 1 TO N
50      READ X
60      PRINT, X
70      T = T * X
```

```
80      NEXT  I
90      G = T   ** (1/N)
100     PRINT
110     PRINT
120     PRINT , G; ` = الوسط الهندسي
130     PRINT
140     PRINT
150     DATA  17, 16, 22, 21, 20, 23, 21,
              19, 15, 13, 23, 17, 20, 29,
              18,22, 16, 25
160     END
```

المخرجات

<u>**البيانات**</u>

```
16
22
21
20
23
21
19
15
13
23
17
20
29
18   الوسط الهندسي 19.662531
22
16
25
```

أما في حالة التوزيعات التكرارية، حيث $س_ر$ X تعني مركز الفئة والتي تكرارها $ك_ر$ F

الفئـــة X	F	مركز الفئة $\overline{X}$ سر	Log $\overline{X}$	كر لوسر F Log $\overline{X}$	الفئة
12.5 – 15.5	2	14	1.14	2.292	
15.5 – 18.5	3	17	1.230	3.690	
18.5 - 21.5	4	20	1.301	5.204	
21.5 – 24.5	5	23	1.362	6.810	
24.5 – 27.5	14	26	1.415	19.810	
27.5 – 30.5	21	29	1.462	30.702	
30.5 – 33.5	59	32	1.505	88.795	
33.5 – 36.5	24	35	1.544	37.056	
36.5 – 39.5	7	38	1.580	11.060	
39.5 – 42.5	6	41	1.613	9.678	
42.5 – 45.5	2	44	1.643	3.286	
45.5 – 48.5	1	47	1.672	1.672	
48.5 – 51.5	2	50	1.699	3.398	
	150			**223.453**	**المجموع**

لو هـ = $\frac{\text{ك}_1 \text{ لو س}_1}{\text{ن}}$

لو هـ = $\frac{223.453}{150}$

لو هـ= 1.4896866

∴ هـ = 30.881

البرنامج التالي يقوم بحساب الوسط الهندسي لبيانات تكرارية:

برنامج لحساب الوسط الهندسي لبيانات متجمعة

```
10    REM
20    DIM   A (13) ,B(13) , C(13) , D(13), E (13), F(13), G(13)
30    F1 = 0   REM  مجموع التكرارات
35    D1  = 0   REM  D  مجموع
40    READN  REM  عدد المجموعات
50    FOR I =  1 TON
```

```
60      READ A (I), B(I),  F(I)  REM
70      F1 = F1 + F(I)        الحد الأدنى،  الحد الأعلى، التكرار
80      C (I) = (A(I) + B(I) ) /2
90      G (I) = LGT (CCI))
100     D (I) = G(I)  *  F (I)
110     D1  =  D1 + D(I)
120     NEXT  I
125     PRINT   ________________________________________
130     PRINT ,               رقم الفئة   الفئة  كر  سر  لوسر  كرلوسر
140     PRINT   ________________________________________
150     PRINT
155     PRINT
160     FOR I = I = 1  TO  N
180     PRINT USING 190, D(I),  G(I), C(I), F(I), B(I) , A(I), I
190     : **.**        *.***    **     **.*   **
200     NEXT   I,
210     PRINT, ______
220     PRINT
230     PRINT  D1; TAB (33); F1;  TAB (50);  المجموع
240     PRINT
250     PRINT
260     H = D1 / F1
270     M = 10 ** H
310     PRINT , M;` =   الوسط الهندسي
320     PRINT
330     PRINT
340     DATA  13, 12, 5, 15.5, 2, 15.5,
350     DATA  18.5 , 3, 18.5, 21.5,  4, 21.5
360     DATA  24.5, 5, 24.5, 27.5, 14, 27.5
              30.5, 21, 30.5, 23.5, 59,
              33., 36.5, 24, 36.5, 39.5,
              7,  39.5,  42.5,  6, 42.5,
              45.5, 2, 45.5, 48.5, 1,
              48.5,  51.5,  2.
400     END
```

المخرجات

رقم الفئة	الفئــة X	كر F	س ر X	لو سر Log X	كرلوسر F Log X
11	15.5 – 12.5	2	14	1.146	2.292
22	18.5 – 15.5	3	17	1.230	3.691
33	21.5 – 18.5	4	20	1.301	5.204
4	24.5 – 21.5	5	23	1.362	6.809
54	27.5 – 24.5	14	26	1.415	19.810
6	30.5 – 27.5	21	29	1.462	30.710
75	33.5 – 30.5	59	32	1.505	88.804
86	36.5 – 33.5	24	35	1.544	37.058
97	39.5 – 36.5	7	38	1.580	11.058
108	42.5 – 39.5	6	41	1.613	9.677
49	45.5 – 42.5	2	44	1.643	3.287
1210	48.5 – 45.5	1	47	1.672	1.672
1311	51.5 – 48.5	2	50	1.699	3.398
12					
18					
المجموع ∑		**150**			**∑ 223.4698**

الوسط الهندسي 30.88858 (من جداول اللوغاريتم المقابل)

2/5/4 خصائص الوسط الهندسي واستخداماته

- عبارة عن قيمة تحويلية Transformed للوسط الحسابي.
  المتوالية الهندسية هي:
  مجموعة القيم المرتبة بحيث تكون النسبة بين كل قيمتين متتاليتين كمية ثابتة.
- الوسط الهندسي هو المقياس الأفضل للنزعة المركزية في حالات الزيادة أو النقصان بنسب ثابتة، كما هو الحال في تقديرات التعداد السكاني، والأسعار.
- الوسط الهندسي هو الأفضل في جميع الحالات التي تستخدم فيها قاعدة الفائدة المركبة لإيجاد الجملة (جـ) التي يؤول إليها مبلغ من المال (أ) بعد (ن) فترة زمنية، بمعدل فائدة ع% عن كل فترة، على النحو التالي:
  جـ = أ (1+ ع)ن
- الوسط الهندسي هو الأفضل لإيجاد متوسط التغير النسبي عند استخدام الأرقام القياسية.

6/4 الوسط التوافقي وتطبيقاته على الحاسوب

الوسط التوافقي Harmonic Mean

إذا كانت س$_1$(x_1)، س$_2$ (x_2)، س$_3$ (x_3) س$_{ن}$(x_n) هي قيم عينية.

فالوسط التوافقي هو:

$$HM = \frac{N \text{ ن}}{\frac{1}{x_1}+\frac{1}{x_2}+\frac{1}{x_3}=....+\frac{1}{x_n}}$$

$$\text{ق} = \frac{\text{ن عدد المتغيرات}}{\sum\left(\frac{1}{\text{س}_{ر}}\right)} \qquad HM = \frac{N}{\sum\left(\frac{1}{X}\right)}$$

فهو إذاً عدد المتغيرات مقسوماً ÷ على مجموع مقلوبات المتغيرات.

مثال:

أوجد الوسط التوافقي للمتغيرات:

17, 23, 13, 15, 19, 21, 23, 20, 21, 22, 16

25, 16, 22, 18, 29, 20

الحل:

$$\text{ق} = \frac{17}{\frac{1}{25}...\frac{1}{21}+\frac{1}{22}+\frac{1}{16}} \qquad Hm = \frac{17}{\frac{1}{16}+\frac{1}{22}+\frac{1}{21}...+\frac{1}{25}}$$

$$\text{ق} = \frac{17}{0.0455+0.0625}$$

$$\text{ق} = \frac{17}{0.8831862}$$

ق = 19.248

فيما يلي برنامج لحساب الوسط التوافقي لقيم عينية باستخدام المعادلة:

حيث: $\frac{N}{T} = HM$

M= الوسط التوافقي
N= عدد المتغيرات
T مجموع مقلوبات المتغيرات

برنامج لحساب الوسط التوافقي لمجموعة مفردات

```
10    REM
20    T = 0   REM   مجموع مقلوب البيانات
30    PRINT     البيانات
40    PRINT ,
50    READ N, REM   عدد الأرقام
60    FOR  I  =  1  TO  N
70    READ   X
80    PRINT   X
90    T = T+1/X
100   NEXT  I
110   PRINT
120   PRINT
130   M = N /T
140   PRINT, M  ;`=  الوسط التوافقي
150   PRINT
160   PRINT
170   17, 16, 22, 21, 20, 23, 21,
      19, 15, 13, 23, 17, 20, 29,'
      18, 22, 16, 25
180   END
```

المخرجات

البيانات

16
22
21
20
23
21
19
15
13
23
17
20
29
18 الوسط التوافقي = 19.24847
22
16
25

أما في حالة البيانات المبوبة في جدول تكراري مكون من ن ، ن(ف) فئة.

فالوسط التوافقي هو:

$$ق = \frac{ك_1 + ك_2 + ك_3 + ... + ك_ف}{\frac{ك_1}{س_1} + \frac{ك_2}{س_2} + \frac{ك_3}{س_3} .... + \frac{ك_ف}{س_ف}}$$

$$\therefore ق = \frac{\sum_{ر=1}^{ف} ك_ر}{\sum \left(\frac{ك_ر}{س_ر}\right)}$$

$$\therefore ق = \frac{ن}{\sum \left(\frac{ك_ر}{س_ر}\right)}$$

$$\therefore \text{ق} = \frac{\text{ن}}{\sum\left(\frac{\text{ك ر}}{\text{س ر}}\right)}$$

مثـال:

أوجد الوسط التوافقي للبيانات الواردة أ دناه، والمبينة أدناه:

رقم الفئة	الفئة	ك ر F	مركز الفئة س ر x	$\frac{\text{ك}_\text{ر}}{\text{س}_\text{ر}}$ $\frac{F}{\bar{X}}$
1	12.5 – 15.5	2	14	.143
2	15.5 – 18.5	3	17	.176
3	18.5 – 21.5	4	20	.200
4	21.5 – 24.5	5	23	.217
5	24.5 – 27.5	14	26	.538
6	27.5 – 30.5	21	29	.724
7	30.5 – 33.5	59	32	1.844
8	33.5 – 36.5	24	35	.686
9	36.5 – 39.5	7	38	.184
10	39.5 – 42.5	6	41	.146
11	42.5 – 45.5	2	44	.045
12	45.5 – 48.5	1	49	.021
13	48.5 – 51.5	2	50	.040
$\sum$		**150**		**4.966**

$$\text{ق} = \frac{150}{4.966} = 30.205 \qquad HM = \frac{\sum F}{\sum \frac{F}{X}}$$

أما في حالة التوزيعات التكرارية، فالبرنامج التـالي يقـوم بحسـاب الوسـط التـوافقي Hasmonic Mean، نستخدم البيانات السابقة وباستخدام المعادلة.

حيث:

$$HM = \frac{F1}{D1}$$

الوسط التوافقي = M

مجموع التكرارات (ن) = F1

مجموع مناسيب التكرارات لمراكز الفئات = D1

برنامج لحساب الوسط التوافقي لبيانات متجمعة

```
10    REM
20    DIM A(13), B(13), C(13),  D(13) ,  F(13)
30    F1 = 0          REM              مجموع التكرارات
35    D1 = 0          REM        D              مجموع
40    READ  N         REM عدد المجموعات
50    FOR  I =  1 TO  N
60    READ  A (I) ,  B(I), F(I)   REM    الحد الأدنى، الحد الأعلى، التكرار
70    F1 =  F1 + F(I)
80    C(I) =  (A(I) + B(I)) /2
90    D (I) = F (1) /  C(I)
100   DI  =  D1 + D(I)
120   NEXT   1
125   PRINT,  ____F/X_____X̄______F_________X
130   PRINT,    ك ر/س ر    س ر    ك ر       الفئة    رقم الفئة
140   PRINT, ____________________________________
150   PRINT,
155   PRINT
160   FOR  I =  1  To   N
170   PRINT  USING  180,  D(I),  C(I) ,   F(I),
      B (I),  A(I),  I
180   **.***    ***    **    **.* - **. ***
190   NEXT  I
200   PRINT
210   PRINT
220   PRINT USING 230  ,  D1 , F1  ****  المجموع
230   **  .  ***
240   PRINT
```

```
250     PRINT
300     M = F1 / D1
310     PRINT , M; ` = الوسط التوافقي
320     PRINT
330     PRINT
340     DATA   13, 12.5,, 15.5, 2, 15.5, 18.5,
350     DATA   3, 18.5, 21.5, 4, 21.5, 24.5,
360     DATA   5, 24.5,  27.5, 14, 27.5,
               30.5,  21, 30.5, 33.5, 59, 33.5,
               36.5, 24, 36.5, 39.5, 7, 39.5, 42.5,
               6, 42.5, 45.5, 2, 45.5, 48.5, 1,
               48.5, 51.5,2,
370     END
```

المخرجات

رقم الفئة	الفئة X	ك ر F	س ر $\bar{X}$	F/ $\bar{X}$ ك ر/س ر
1	15.5 – 12.5	2	14	0.143
2	18.5 – 15.5	3	17	0.176
3	21.5 – 18.5	4	20	0.200
4	24.5 – 21.5	5	23	0.217
5	27.5 – 24.5	14	26	0.538
6	30.5 – 27.5	21	29	0.724
7	33.5 – 30.5	59	32	1.844
8	36.5 – 33.5	24	35	0.686
9	39.5 – 26.5	7	38	0.184
1	42.5 – 39.5	6	41	0.146
1	45.5 – 42.5	2	44	0.045
1	48.5 – 45.5	1	47	0.021
1	51.5 – 48.5	2	50	0.040
	المجموع	**150**		

الوسط التوافقي = 30.205

Σ

خصائص الوسط التوافقي واستخداماته:

- الوسط التوافقي هو مقلوب الوسط الحسابي لمقلوبات القيم.
- الوسط التوافقي هو قيمة تحويلية للوسط الحسابي يستخدم بـدلاً منـه في حـالات خاصـة جـداً شأنه في ذلك شأن الوسط الهندسي.

1/6/4 العلاقة بين المتوسط الحسابي، والوسيط والمنوال:

1. إذا كان التوزيع التكراري، قريب من التماثل يعني أنه قريب من الشكل الناقوس، فإن:

   المتوسط الحسابي – المنوال   3   (المتوسط الحسابي – الوسيط)

2. أما إذا كان التوزيع التكراري متماثلاً أي ناقوس الشكل، فإن:

   المتوسط الحسابي = الوسيط = المنوال

ويلاحظ أن المتوسط الحسابي، يمكن حسابه بسهولة، ويأخذ بعين الاعتبار جميع القيم الواردة في التوزيع، ويتأثر بها. وبالتالي فإنه يعتبر أحسن مقاييس النزعة المركزيـة، إلا أنـه يعـاب عليـه أنـه يتـأثر بالقيم المتطرفة أو الشاذة، ومن هذه الحالة يفضل استخدام الوسيط أو المنوال...

كذلك لا يصلح المتوسط الحسابي في حساب البيانات التي تحتـوي عـلى فئـات مفتوحـة، ولـذلك يفضل استخدام الوسيط في هذه الحالة، وذلك لأنه يعتمد في حسابه على التكرارات وليس عـلى مراكـز الفئات.

ويمكن أيضاً استعمال المنوال في الفئات المفتوحة، إلا أنـه يعـاب عـلى الوسـيط أنـه لا يعتمـد في حسابه على كل القيم الواردة في التوزيع وكذلك لا يصلح ، لا يصلح في إعطاء فكرة صحيحة عن النزعة المركزية، عندما تكون البيانات متجمعـة في فئـات متباعـدة، ولـذلك يستحسـن اسـتعمال المنـوال لأنـه أفضل من الوسيط والمتوسط الحسابي في هذه الحالة.

الوحدة الخامسة

مقاييس التشتت

MEASURFS OF DISPERSION

إن مقاييس النزعة المركزية ليست كافية وحدها، وذلك لأنه قد يساوي متوسطات في مجتمعين ولكن يختلف في صفة أخرى، مثل قد تتساوى الولايات المتحدة وقطر في متوسط دخل الفرد، إلا أن توزيع الدخل قد يكون مختلفاً في هاتين الدولتين، ومن هنا جاءت الحاجة إلى ما يسمى بمقاييس التشتت.

مثـال:

مصنع X خير بين استخدام نوعين من الآلات (آلة أ وآلة ب)، وأراد المصنع اختيار درجة جودة الإنتاج لهاتين الآلتين، فأنتجت (9) وحدات من كل من الآلتين. وفحصت السلعة المنتجة وقدرت درجة الجودة الإنتاج، فكانت على النحو التالي:

درجة جودة الإنتاج باستخدام الآلة أ	درجة جودة الإنتاج باستخدام الآلة ب
90	126
116	128
198	104
24	114
170	136
92	118
70	118
144	100
64	64
1008	1008

ولو استعمل المصنع X المتوسط الحسابي كمعيار للمقارنة، فإنه يقيس متوسط درجة جودة الإنتاج.

متوسط درجة الإنتاج للآلة أ $= \frac{1008}{9} = 112$

متوسط درجة الإنتاج للآلة ب $= \frac{1008}{9} = 112$

فمن حيث المتوسط، يلاحظ بأن الآلتان متساويتان، وبالتالي فإنه لا يمكن المفاضلة بينهما على هذا النحو.

بينما أرقام جودة الإنتاج تبين اختلاف في الجودة، فالوحدات المنتجة باستخدام الآلة (أ) أكثر تبايناً من الوحدات المنتجة باستخدام الآلة (ب). فالآلة ب إذن أكثر استقراراً في الإنتاج، وبالتالي فهي الأفضل.

وهذا المثال يبين بوضوح مدى أهمية مقاييس التشتت بجانب مقاييس النزعة المركزية.

1/5 أهم مقاييس التشتت:

- المدى المطلق Absolute Range
- الانحراف المتوسط Average Deviation
- والانحراف الربيعي.
- الانحراف المعياري.

1- المدى المطلق:

هو الفرق بين أعلى قيمة وأصغر قيمة.

المدى المطلق = أعلى قيمة – أصغر قيمة

المدى المطلق في درجة الإنتا باستخدام الآلة، أ=

198 – 24 = 174

المدى المطلق في درجة جودة الإنتاج باستخدام الآلة ب=

136 – 64 = 72

ومن هذه النتيجة يمكن القول بأن هناك تفاوت أو تباين بين جودة الوحدات المنتجة باستخدام الآلة (أ) أكبر منها باستخدام الآلة (ب)، ومن ثم يمكننا اختيار الآلة (ب).

أما إذا كانت قيم مبوبة في فئات فإن المدى المطلق، يمكنحسابه، عن طريق أخذ الفرق بين الحد الأعلى للفئة الأخيرة، والحد الأدنى للفئة الأولى في التوزيع، كما هو مبين من جدول التوزيع التكراري.

الفئات	F
20 – 24	5
25 – 29	11
30 – 34	39
35 – 39	24
40 – 44	10
45 – 49	6
المجموع	**95**

وحيث أن الحد الأدنى للفئة الأولى يساوي 20 والحد الأعلى للفئة الأخيرة في الجدول يساوي 49.

فإن المدى المطلق يساوي 29= 20 – 49

وبالرغم من سهولة المدى المطلق إلا أنه مضلل وغير دقيق، لا يمكن حسابه في حالـة التوزيعـات التكرارية المفتوحة.

2/5 الانحـــراف الربيعي

الانحراف الربيعي هو عبارة عن قيمة الربيع الأعلى – ناقصاً قيمة الربيع الأدنى مقسـوماً ÷ عـلى 2 أي:

$$\text{الانحراف الربيعي} = \frac{\text{الربيع الأعلى} - \text{الربيع الأدنى}}{2}$$

ويلاحظ أنه كلما اتسع المدى بين الربيع الأعلى والربيع الأدنى كلما دلنا ذلك على تشتت مفردات الظاهرة المراد دراستها، مثال لاحتساب الانحراف الربيعي.

الفئات	التكرار F	التكرار المتجمع الصاعد
أقل من 12.5	20	20
أقل من 15.5	64	84
أقل من 18.5	156	240
أقل من 21.5	69	336
أقل من 24.5	40	376
أقل من 27.5	24	400
	400	

نحسب أولاً الربيع الأدنى

ترتيب الربيع الأدنى $\frac{N}{4} = \frac{400}{4} = 100$

فالفئة التي تحوي الربيع الأدنى هي 18-16 (أقل من 18.5 وتكرارها (156) وطول الفئة 2.

$$Q_1 = L + \left(\frac{P}{F}\right)^1 .R$$

$$= 15.5 \frac{\left(\frac{400}{4} - 84\right)^2}{156}$$

$$= 15.5 + \frac{32}{156} = 16$$

ثم بعد ذلك نقوم بحساب الربيع الأعلى بنفس الطريقة

$$Q_2 = L + \left(\frac{P}{F}\right)^1 .R$$

$$= 18.5 \frac{\left(\frac{3 \times 400}{4} - 240\right)^{.2}}{96}$$

$$= 18.5 + \frac{120}{96} = 18.5 + 1.2 = 19.7$$

الانحراف الربيعي $\frac{19.7 - 16}{2}$

$$= \frac{3.7}{2} = 1.8$$

ويمكـن اسـتخدام الانحـراف الربيعـي في حالـة الجـداول المفتوحـة والتـي يصـعب فيهـا إيجـاد الانحراف المعياري، والانحراف المتوسط، كما أنه لا يتأثر بالقيم المتطرفة، ولكنه يعتمد في حسـابه عـلى قيمتين فقط، ويهمل بقية القيم.

3/5 الإنحراف المتوسط **Average deviation**

لإيجاد الانحراف المتوسط، يحسب أولاً المتوسط الحسابي، ثم توجد انحراف كـل مركـز فئـة عـن المتوسط الحسابي، مع إهمال الإشارة، ثم ضرب كل انحراف في التكرار المقابل له لنحصل على مجموعها $\frac{\sum 1\overline{X} - \overline{X}1F}{N}$.

وبالتالي فإن الانحراف المتوسط A.D يكون:

$$A.D = \frac{\sum df}{N}$$

حيث d عبارة عن الانحرافات $d = 1X - \overline{X}1$

إذا كانت لدينا قيم مفردة، فإنه يمكن إيجاد الانحراف المتوسط لمثل هذه البيانات، كالمثال الآتي: 4,6,12,16,22,

توجد أولاً: المتوسط الحسابي $\overline{X} = \frac{4+6+12+16+22}{5} = 12$

وبذلك يكون الانحراف المتوسط:

$$A.D = \frac{(4-12)+(6-12)+(12-12)+(16-12)+(22-12)}{5}$$

$= 5.6$

أما في حالة البيانات المبوبة، فيمكن إيجاد الانحراف المتوسط وذلك باستخدام التوزيع التالي:

الفئات	F	X	FX	$d= \lvert X - \overline{X} \rvert$	FD
60 - 62	10	61	610	6.45	68.50
63 – 65	36	64	2304	3.45	124.20
66 – 68	44	67	5628	.45	37.80
69 – 71	54	70	3780	2.55	137.70
72 – 74	16	73	1168	5.55	88.80
	460		**13490**		**453.00**

أولاً: توجد المتوسط الحسابي

$$\overline{X} = \frac{13490}{200} = 67.45$$

ثم بعد ذلك توجد الانحراف المتوسط:

$$A.D = \frac{\sum df}{N}$$

$$= \frac{453}{200} = 2.265$$

وخواص الانحراف المتوسط:

1. لا يعتمد في حسابه على جميع القيم.
2. لا يمكن حسابه في حالة التوزيعات التكرارية المفتوحة.
3. يتأثر بالقيم المتطرفة لأن انحرافها عن الوسط الحسابي تكون كبيرة.

4/5 الانحراف المعيـــاري: **Standard Deviation**

وهو من أهم مقاييس التشتت، ويستعمل كثيراً بالرغم من صعوبة حسابه. ويمكن حساب الانحراف المعياري عن بيانات غير مبوبة، فلو وجدت مشاهدات تأخذ الصيغة التالية: 9, 6, 5, 11, 1, 6, 7, 3 فإنه يمكن حساب الانحراف المعياري لمثل هذه البيانات.

فمجموع المشاهدات = 48.

أول خطوة في إيجاد الانحراف المعياري لهذه البيانات هو إيجاد المتوسط الحسابي لها= [مجموع عدد المشاهدات÷ 8]

∴ المتوسط الحسابي يكون = $\overline{X} = \frac{48}{8} 6$

ثم نجد الانحرافات عن المتوسط الحسابي، ونجد بعد ذلك مربع الانحرافات عن المتوسط الحسابي، وذلك على النحو التالي:

$D = X - \overline{X}$		d_2
9- 6	3	9
6 – 6	-	-
5 – 6	-1	1
11 – 6	5	25
1 – 6	-5	25
6 – 6	-	-
7 – 6	1	1
3 – 6	-3	9
	المجموع	70

وإذا رمزنا للانحراف المعياري بالرمز (σ)

$$\sigma = \sqrt{\frac{\sum d2}{N}} = \sqrt{\frac{70}{8}} = \sqrt{8.75}$$

$\sigma = 2.95$

5/5 التبايـــن: **Variance**

ويمكن إيجاد ما يسمى بالتباين Variance وهو المقيـاس الرقمـي الـذي يصـلح كمقيـاس لدرجـة تشتت القيم.

والتباين عبارة عن مجموع انحرافات القيم ÷ على عددها.

ويلاحظ بأن هذه الطريقة في إيجاد الانحراف المعياري تأخذ جهداً كبيراً في العمليات الحسابية.

ولتبسيط العمليات الحسابية يمكن حساب الانحراف المعياري بالقانون التالي:

$$\sigma = \sqrt{\frac{\sum X^2}{N} - \left(\frac{\sum X}{\boldsymbol{N}}\right)^2}$$

ويمكن حسابه باستعمال المثال السابق:

X	X^2
9	81
6	36
5	25
11	121
1	1
6	36
7	49
3	9
48	358

وتطبيق المعادلة السابقة، يكون الانحراف المعياري، كالآتي:

$$\sigma = \sqrt{\frac{\sum X^2}{N} - \left(\frac{\sum X}{N}\right)^2}$$

$$= \sqrt{\frac{358}{8} - \left(\frac{48}{8}\right)^2} = 2.95$$

والتباين هو عبارة عن:

$$\sigma = \frac{\sum X^2}{N} - \left(\frac{\sum X}{N}\right)^2$$

(مربع الانحراف المعياري)

ويمكن حساب الانحراف المعياري لقيم مبوبة، كما هو مبين في الجدول التالي:

الفئات	F	X	FX	X^2	FX^2
48 – 52	2	50	100	2500	5000
53 – 57	3	55	165	3025	9050
58 – 62	4	60	240	3600	1440
63 – 67	6	65	390	4225	25350
68 – 72	3	70	210	4900	14700
73 - 77	2	75	150	5625	11250
المجموع	**20**		**1255**		**79775**

$$\sigma = \sqrt{\frac{79775}{20} - \left(\frac{1255}{20}\right)^2}$$

$$= \sqrt{8988.75 - (62.75)^2}$$

$$= \sqrt{8988.75 - 3937.56} \quad = \quad \sqrt{51.19}$$

$$= \boxed{7.15}$$

ويلاحظ بأن هذه الطريقة مطولة، وتحتاج إلى الكثير من العمليات الحسابية، إلا أنه يمكن اختصار هذه العمليات الحسابية باستعمال طريقة مختصرة، حيث يمكن أخذ الانحرافات عن وسط فرضي A=60 ، وذلك على النحو التالي:

الفئات	F	X	D = X-A	D^2	DF	D^2F
48 – 52	2	50	-10	100	-20	200
53 – 57	3	55	-5	25	-15	75
58 – 62	4	60	d = X -A d =60-60=0	-	-	-
63 – 67	6	65	5	25	30	150
68 – 72	3	70	10	100	30	300
73 - 77	2	75	15	225	30	450
المجموع	**20**				**55**	**1175**

$$\sigma = \sqrt{\frac{\sum Fd^2}{N} - \left(\frac{\sum Fd}{N}\right)^2}$$

$$= \sqrt{\frac{1175}{20} - \left(\frac{55}{20}\right)^2}$$

$$= \sqrt{58.750 - 7.5625}$$

$$= \boxed{7.15}$$

k = طول الفئة = 5

الوسط الفرضي = A=60

ويمكن إيجاد الانحراف المعياري بطريقة مختزلة، وذلك بتسجيل العمليات الحسابية، بصورة أكثر لتستخدم نفس المثال:

الفئات	F	X	d	$d^1 = \frac{d}{k}$	d_1^2	Fd_1^2	Fd_1
48 – 52	2	50	-10	-2	4	8	-4
53 – 57	3	55	-5	-1	1	3	-3
58 – 62	4	60	-	-	-	-	-
63 – 67	6	65	5	1	1	6	6
68 – 72	3	70	10	2	4	12	6
73 - 77	2	75	15	3	9	18	6
المجموع	**20**					**47**	**11**

ويمكن إيجاز الخطوات السابقة:

- قمنا باختيار وسط فرضي = 60.
- ثم حسبنا انحراف مركز الفئة X عن الوسط الفرضي.
- ثم أخذنا عامل مشترك، وقسمنا ÷ الانحراف (d) على العامل المشترك (K) لينتج لنا (d1).
- ثم وجدنا مجموع الانحراف Xت مضروباً في التكرار، ومجموع مربع الانحراف X مضروباً في التكرار.
- ثم حسبنا الانحراف المعياري، كالآتي:

$$\sigma = \sqrt{\frac{\sum Fd_1^2}{N} - \left(\frac{\sum Fd_1}{N}\right)^2}$$

$$= 5\sqrt{\frac{47}{20} - \left(\frac{11}{20}\right)^2}$$

$$= 5\sqrt{2.35 - (.55)^2}$$

$$= 5\sqrt{2.05} \quad \cong \quad \text{مقربة}$$

$$= 5 \times 1.43 = \boxed{7.15}$$

6/5 معامل الاختــلاف **Coefficient Variation**

عند مقارنة التشتت بالنسبة للمشاهدات لمتغير واحد على مستويات مختلفة، أو مشاهدات لمتغيرات مختلفة، يجب أن نحول مقياس التشتت لهذه المشاهدات إلى مقياس نسبي، وذلك حتى يمكن التخلص من تأثير وحدات القياس.

ولعمل هذا نستخدم مقياساً يسمى بمعامل الاختلاف ويعرف على أنه نسبة الانحراف المعياري ÷ إلى المتوسط الحسابي.

$$\text{معامل الاختلاف} = \frac{\text{الانحراف المعياري}}{\text{المتوسط الحسابي}} \times 100$$

فإذا فرضنا أن الانحراف المعياري للأعمال في عينة من الموظفين هو (16) سنة، والمتوسط الحسابي لأعمارهم هو (80) سنة، بينما كان الانحراف المعياري لأعمال عينه من الطلبة هو (12) سنة، والمتوسط الحسابي لأعمارهم (40) سنة.

فيمكن حساب معامل الاختلاف للتفاوت في السن، بين العينين، وذلك كالآتي:

$$\text{معامل الاختلاف لأعمار الموظفين} = \frac{16}{80} \times 100$$

$$= \%20$$

$$\text{معامل الاختلاف لأعمال الطلبة} = \frac{12}{40} \times 100$$

$$= \%30$$

ومن هذه النتائج يمكننا القول بأن أفراد عينة الطلبة أكثر تبايناً من حيث العمر من عينة الموظفين.

7/5 تطبيقات مقاييس التشتت على الحاسب الإلكتروني Moments

مقاييس التشتت والعزوم Measures of Dispersion & Moment

أهم مقاييس التشتت:

1. المدى Range
2. الانحراف الربيعي Quartile Deviation
3. الانحراف المتوسط Mean Deviation

الانحراف المعياري Standard Deviation

- **المــدى Range**

هو الفرق بين أكبر قيمة وأصغر قيمة. وفي حالة البيانات المبوبة في فئات

- يكون المدى = هو الفرق بين الحد الأدنى للفئة الأولى، والحد الأعلى للفئـة الأخـيرة (العليـا) في حالة البيانات المبوبة.
- وقد يكون الفرق أيضاً بين مركز الفئة الأخيرة ومركز الفئة الأولى.
- وقد يكون المدى هو الفرق بين الربيع الثالث (الأعلى) (,75%)والربيع الأول (,25%)، وهو ما يسمى "بالمدى الربيعي".

- **الانحراف الربيعي Quartile Deviation**

وهو نصف المدى الربيعي، وبذلك يكون الانحراف الربيعي

$$= \frac{(\%25,) - (\%75,)}{2}$$

يعتبر الانحراف الربيعـي مفيـداً جـداً في حالـة التوزيعـات ذات الفئـات المفتوحـة والتـي يمكـن استخراج انحرافاتها الربيعية دون المقاييس الأخرى.

مثال: أوجد المدى، والانحراف الربيعي للبيانات التالية:

رقم الفئة	الفئات	F كر التكرار	$\overline{X}$ سر مركز الفئة	المتجمع الصاعد
1	15.5 – 12.5	2	14	2
2	18.5 – 15.5	3	17	5
3	21.5 – 18.5	4	20	9
4	24.5 – 21.5	5	23	14
5	27.5 – 24.5	14	26	28
6	30.55 – 27.5	21	29	49
7	33.5 – 30.5	59	32	108
8	36.5 – 33.5	24	35	132
9	39.5 – 36.5	7	38	139
10	42.5 – 39.5	6	41	145
11	45.5 – 42.5	2	44	147
12	48.5 – 45.5	1	47	148
13	51.5 – 48.5	2	50	150
	المجموع	**150**		$\sum$

المدى = الحد الأعلى للفئة الأخيرى (العليا) – الحد الأدنى للفئة الأولى

= 51.5 - 12.5 = 39.0

وباعتبار المدى هو الفرق بين مركزي الفئتين السالفتين:

المدى = 50-14 = 36

البرنامج التالي يقوم بحساب المدى لبيانات متجمعة، والبيانات المستخدمة هي نفسها بالمثال السابق:

برنامج لحساب المدى لبيانات متجمعة

```
10    REM
20    DIM A(13), B(13), F(13), C(13)
30    T=0
40    READ  N  REM  عدد القيم
50                          رقم الفئة    الفئة    مركز الفئة
60    ________________________________________________
70    ***          **.* - **.*     **
75    PRINT USING   60
80    PRINT USING   50
90    PRINT USING   60
100   PRINT
```

```
105     FOR  I = 1 TO N
110     READ A(I), B(I), F(I),  REM  الحد الأدنى، الحد الأعلى، التكرار
120     C(I) – (A(I) + B(I)) /2  REM مركز الفئة
140     PRINT USINT 70, C(I) , B(I), A(I) , I
150     NEXT  I
160     R1 = B(N) – A(1)
170     R2 = C(N) – C(1)
180     PRINT , R1 ;` =   المدى 1`
190     PRINT
200     RINT, R2;` =        المدى2`
210     PRINT
220     PRINT
230     DATA  13, 12.5, 15.5, 2, 15.5, 18.5, 3, 18.5, 21.5,
        4, 21.5, 24.5, 5, 24.5, 27.5, 14, 27.5, 30.5,
        5, 33.5, 59, 33.5, 36.5, 24, 36.5, 39.5, 7,39.5, 42.5,
        6, 42.5, 45.5, 2, 45.5, 48.5, 1, 48.5, 51.5, 2.
260     END
```

المخرجات

رقم الفئة	الفئات	مركز الفئة
1	15.5 -12.5	14
2	18.5 – 15.5	17
3	21.5 – 18.5	20
4	24.5 – 21.5	23
5	27.5 – 24.5	26
6	30.5 – 27.5	29
7	33.5 – 30.5	32
8	36.5 – 33.5	35
9	39.5 – 36.5	38
10	42.5 – 39.5	31
11	45.5 – 42.5	44
12	48.5 – 45.5	47
13	51.5 – 48.5	50

المدى 1= 39

المدى 2= 36

- **الانحراف الربيعي:**

$$\frac{\text{الربيع الأعلى - الربيع الأدنى}}{2}$$

وفيما يلي برنامج حساب الانحراف الربيعي للبيانات الواردة في نفس المثال والمثال السابق، علماً بأن معاملة الانحراف الربيعي المستخدمة هي:

$$y = \frac{Q_1 - Q_3}{2}$$

حيث:

الانحراف الربيعي = Y

الربيع الأعلى = Q_1

الربيع الأدنى = Q_3

كما أن :
$$\boldsymbol{Q}(1) = \boldsymbol{l}(1) + \frac{\boldsymbol{S}(1) - 0(1)\boldsymbol{P}(1)}{\boldsymbol{W}(1)}$$

برنامج لحساب الانحراف الربيعي لبيانات متجمعة

```
10    REM
20    DIM  A(13), B(13), C(13), D(13), E (13), F(13) G(13)
30    T = 0
40    READ  N
50    FOR I = 1  TO  N
60    READ A(I), B(I), F(I) REM  الحد الأعلى، الحد الأدنى، والتكرار
70    T  =  T +F (I)                REM  مجموع التكرارات
80    NEXT I
90    C(I) = F(I)
100   FOR 1 = 2 TO N
110   C (I) = C (I-1) + F (I)
120   NEXT   I
130   رقم الفئة     الفئة    التكرار   المتجمع التكراري الصاعد
140   ________________________________________
150   ** *      **       **.*  -  **.*    **
160   ________________________________________
```

```
170        ***        المجموع
180
185        PRINT  USING          150
190        PRINT  USINT          130
200        PRINT USING           140
210        PRINT USING           150
220        PRINT
230        FOR  I = 1  TO N
240        PRINT USING 160, C(I), F(I), B(I), B(I), A(I), I
250        NEXT  I
260        PRINT USING 170
270        PRINT USING 180, T
280        PRINT
290        S1 = TX 3/4
300        S3  =  T/4
310        FOR I = 1  TO  N
320        IF C(I) > S1  THE 340
330        J = I
340        NEXT  I
350        FOR I = 1  TO  N
360        IF C(I) > S3   THEN  380
370        V  =  I
380        NEXT  I
390        PRINT USING 430, B(J+1), A(J+1), J+1
400        PRINT
410        PRINT USING 440, B(V+1), A(V+1), V+1
420        PRINT
430        فئة الربيع الأعلى هي الفئة رقم  ** وهي  *.**-*.**
440        فئة الربيع الأدنى هي الفئة رقم  ** وهي *.** - *.**
450        GO  SUB  610
460        Q1 = FNA (L1, S1, 01, P1, W1)
470        PRINT,  Q1;  `  =  الربيع الأعلى
480        PRINT
510        GO SUB 660
520        Q3 =  FNA (L3, S3, 03, P3, W3)
530        PRINT, Q3; ` =  الربيع الأدنى
550        PRINT
560        Y = (Q1-Q3)/2
```

```
570     PRINT, Y;` =   الانحراف الربيعي
580     PRINT
590     PRINT
600     STOP
610     L1 = A (J+1)
620     O1 = C(J)
630     P1 = B(J) – A(J)
640     W1 = F (J+1)
650     RETURN
660     L3 =  A (v+1)
670     O3 = C(V)
680     P3 =  B(V) – A (V)
690     W3 = F (V+1)
700     RETURN
710     DEF FNA (L, X,O, P, W)= L+((X-0)*P)/W
720     DATA 13,12,5,15.5,2,15.5,18.5,3,18.5,
        DATA 21.5, 4, 21.5, 24.5, 5, 24.5, 27.5, 14,
        DATA  27.5, 30.5, 21, 30.5, 33.5, 59, 33.5,
        DATA 36.5, 24, 36.5, 39.5, 7, 39.5, 42.5, 6,
        DATA 42.5, 45.5, 2, 45.5, 48.5, 1, 48.5, 51.5, 2,
750     END
```

المخرجات

رقم الفئة	الفئات	التكرار	المتجمع التكراري الصاعد
1	15.5 -12.5	2	2
2	18.5 – 15.5	3	5
3	21.5 – 18.5	4	9
4	24.5 – 21.5	5	14
5	27.5 – 24.5	14	28
6	30.5 – 27.5	21	49
7	33.5 – 30.5	59	108
8	36.5 – 33.5	24	132
9	39.5 – 36.5	7	139
10	42.5 – 39.5	6	145
11	45.5 – 42.5	2	147
12	48.5 – 45.5	1	148
13	51.5 – 48.5	2	150
	المجموع	**150**	$\sum$

- فئة الربيع الأعلى هي الفئة رقم 8 وهي 33.5 – 36.5
- فئة الربيع الأدنى هي الفئة رقم 6 وهي 27.5 – 30.5

الربيع الأعلى = 34.0625

الربيع الأدنى = 28.85713

الانحراف الربيعي = 2.602684

- **الانحراف المتوسط Average Deviation**

هو مجموع القيم المطلقة للانحرافات عن الوسط الحسابي ÷ مقسوماً ÷ على عددها، والقيمة المطلقة Absolute value هي:

$$|س| = \begin{cases} س & \text{إذا كانت س } \leq \text{صفراً} \\ -س & \text{إذا كانت س } \geq \text{صفراً} \end{cases}$$

وهذا معناه إهمال الإشارة السالبة، واعتبار الانحراف يمثل بعداً لا يهم اتجاهه، بذلك يكون:

$$\text{الانحراف المتوسط} = \frac{\sum_{ر=1}^{ن} |س_ر - \overline{س}|}{ن}$$

خطوات إيجاد الانحراف المتوسط:

1. إيجاد المتوسط الحسابي.
2. إيجاد انحراف كل مركز فئة عن المتوسط الحسابي (مع إهمال الإشارة).
3. ضرب كل انحراف x التكرار المقابل له لنحصل على مجموعها.

مثال: أوجد الانحراف المتوسط على القيم أدناه:

4, 8, 6, 3, 1, 2

$$\text{المتوسط الحسابي} = \overline{X} = \frac{4+8+6+3+1+2}{6} = \frac{24}{6} = 4$$

الانحراف المتوسط =

$$A.D = \frac{(4-4)+(8-4)+(6-4)+(3-4)+(1-4)+(2-4)}{6}$$

الانحراف المتوسط= $A.D = \frac{0+4+2+1+3+2}{6} = \frac{12}{6} = 2$

مثال: أوجد الانحراف المتوسط للبيانات التالية:

رقم الفئة	الفئة	ك ر F	مركز الفئة س ر X	$D = X - \bar{X}$	FD |ر - $\bar{س}$| ك ر	FX |س| س ر×ك ر
1	12.5 – 15.5	2	14	= 17.48-	34.96	28
2	15.5 – 18.5	3	17	= 14.48 –	43.44	51
3	18.5 – 21.5	4	20	= 11.48-	45.92	80
4	21.5 – 24.5	5	23	= 8.48 -	42.40	115
5	24.5 – 27.5	14	26	= 5.48-	76.72	364
6	27.5 – 30.5	21	29	= 2.48-	52.08	609
7	30.5 – 33.5	59	32	= 0.52	30.68	1888
8	33.5 – 36.5	24	35	= 3.52	84.48	840
9	36.5 – 39.5	7	38	= 6.52	45.64	266
10	39.5 – 42.5	6	41	= 9.52	57.12	246
11	42.5 – 45.5	2	44	= 12.52	25.04	88
12	45.5 – 48.5	1	47	= 15.52	15.52	47
13	48.5 – 51.5	2	50	= 18.52	37.04	100
$\sum$	$\sum$ F =	**150**		$\sum$ F =	**591.04**	**4722**

المتوسط الحسابي $\sum \frac{FX}{F} = \frac{4722}{150} = 31.48$

الانحراف المتوسط $A.D = \frac{\sum Ed}{F} = \frac{591.04}{150} = 3.94 == 3.94$

- **الانحراف المعياري: Standard Derivation**

للتخلص من الإشارات السالبة للانحرافات عن الوسط الحسابي، وهـي بتربيـع تلـك الانحرافـات لتصـبح جميعهـا موجبة، ويكـون مجموعهـا $\sum$(س ر -$\bar{س}$)2 يسـمى متوسـط مجمـوع مربعـات الانحرافـات المـذكورة أعـلاه بالتبـاين Variance، وقد عدّل لاعتبارات

الاستدلال الإحصائي تعديلاً طفيفاً ليصبح (ن-1) بدلاً من (2) وبذلك يكون تباين مفردات العينة هو:

$$ع^2 = \frac{\sum_{ر=1}^{ن} (س_ر - \bar{س})^2}{(ن - 1)}$$

$$\text{الانحراف المعياري } ع = \sqrt{\frac{\sum (س_ر - \bar{س})^2}{(ن - 1)}}$$

مثال: أوجد التباين، والانحراف المعياري للقيم

25, 16, 22, 18, 29, 20, 17, 23, 13, 15, 19, 21, 23, 20, 21, 22, 16,

الحل:

$\bar{س} = 20$

$ن = 17$

$$ع^2 = \sum (س_ر - \bar{س})^2 = (-4)^2 + (2)^2 + (1)^2 + \ldots + \ldots + (-4)^2 + (5)^2$$

$$= 16 + 4 + 1 \qquad + 16 + 25$$

$$= 245$$

$$\frac{\sum (س_ر - \bar{س})^2}{(ن - 1)} = \frac{254}{16} = 15.875 \quad ع^2 \text{ التباين}$$

$$ع = \sqrt{15.875} = 3.984$$

$$ع^2 = \frac{\sum س_ر^2 - \frac{(\sum س_ر)^2}{2}}{(ن - 1)}$$

فيما يلي برنامج لحساب التباين والانحراف المعياري للقيم أعلاه، علماً بأن المعادلة المستخدمة هنا هي:

$$V = \frac{T}{(N - 1)}$$

$$R = \sqrt{V}$$

حيث:

التباين =V

مجموع مربعات الانحراف = T

عدد المتغيرات =

وبالتالي فإن:

$$R = \sqrt{V}$$

حيث R الانحراف المعياري

برنامج لحساب التباين والانحراف المعياري لمجموعة مفردات

```
10     REM
20     DIM  X  (17)
30     S = 0
40     READ N REM  عدد القيم
50     FOR  I =   1  TO N
60     READ  X  (I)
70     S = S + (I)
80     M =  S/N
100    PRINT USING 300
105    PRINT USING 290
110    PRINT USING 300
120    PRINT
130    FOR  I  =1  TO  N
```

```
140     D =  X(I) – M
150     T =  T + D XX2
160     PRINT USING 310, X(1), D,D**2
170     NEXT  I
180     PRINT
190     PRINT USING 320
200     PRINT USING 330 , T
220     PRINT
230     V = T/ (N-1)
240     R = SOR (V)
250     PRINT, V;` = VARIANCE, التباين
255     PRINT
260     PRINT, R; ` = STAQNDARD DEVIATION الانحراف المعياري
270     PRINT
280     PRINT
290     القيمة س        س-س-  (س-س-)2
300     ______________________________________________
310     ***              **                    **
320     _____________________________________________
330     ****
340     DATA  17,16,22,21,20,2,23,21,19,
                15,13, 23, 17, 20, 29, 18, 22,
                16, 25
350     END
```

المخرجات

القيمة س	س-س-	(س-س-)^2
16	-4	16
4	2	22
1	1	21
0	0	20
9	3	23
1	1	21
1	-1	19
25	-5	15
49	-7	13
9	3	23
9	-3	17
0	0	20
81	9	29
4	-2	18
4	2	22
16	-4	16
25	5	25
المجموع		**254**

التباين = 15.875

الانحراف المعياري = 3.984344

أما في حالة التزيعات التكرارية، فترجع مربعات الانحرافات بتكراراتها ليصبح على النحو التالي:

$$ع^2 = \frac{\sum_{ر=1} (س_ر - \bar{س})^2 ك_ر}{(ن - 1)}$$

وهذه أيضاً يمكن تعديلها ليكون التباين:

$$ع^2 = \frac{\sum س_ر^2 ك_ر - \left(\frac{\sum س_ر ك_ر}{ن}\right)^2}{ن - 1}$$

رقم الفئة	الفئات	F $ك_ر$	$\overline{X}$ $س_ر$	$F\overline{X}$ $س_ر ك_ر$	$\overline{X}^2$ $س_ر^2$	$س_ر^2 ك_ر$
1	15.5 – 12.5	2	14	28	196	392
2	18.5 – 15.5	3	17	51	289	867
3	21.5 – 18.5	4	20	80	400	1600
4	24.5 – 21.5	5	23	115	429	2645
5	27.5 – 24.5	14	26	364	676	9464
6	30.5 – 27.5	21	29	609	841	17661
7	33.5 – 30.5	59	32	1888	1024	60416
8	36.5 – 33.5	24	35	840	1225	29400
9	39.5 – 36.5	7	38	266	1444	10108
10	42.5 – 39.5	6	41	246	1681	10086
11	45.5 – 42.5	2	44	88	1936	3872
12	48.5 – 45.5	1	47	47	2209	2209
13	51.5 – 48.5	2	50	100	2500	5000
$\sum$	**المجموع**	**150**		**4722**		**153720**

$$ع^2 = \frac{153720 - \frac{4722 \times 4722}{150}}{149}$$

$\therefore$ $ع^2 = 34.03651$

$ع = 5.834$

ملاحظة: قيمة التباين دائماً تكون موجبة لأنها مجموع مربعات، لذلك فإن أقل قيمة للتباين = صفر، وتعني التجانس التام.

نستخدم البرنامج التالي لحساب الانحراف المعياري لبيانات مبوبة، وباستخدام المعادلة:

$$V = \frac{(T3 - (T2)^2}{T1 - 1}$$

حيث:

التباين V=

مجموع مربعات مراكز الفئات مرجحة بتكراراتها T_3=

مجموع مراكز الفئات مرجحة بتكراراتها T2 =

مجموع التكرارات T1 =

وكذلك: R = $\sqrt{V}$

الانحراف المعياري R =

برنامج لحساب التباين والانحراف المعياري لبيانات مبوبة

```
10     REM
20     DIM A(13), B(13), C(13), D(13), E(13), F(13), G(13)
30     F1 = 0
40     READ   REM  عدد المشاهدات
50     FOR I = 1  TO  N
60     READ A(I), B(I), F(I)  REM   الحد الأدنى، الحد الأعلى ، التكرار
70     NEXT  I
80     PRINT USING 320
85     PRINT USING 300
90     PRINT USING 310
100    PRINT USING 320
110    PRINT
120    FOR  I =  1  TO N
130    C(I) = (A(I) +B(I))/2
140    C(I) = C(I) * F(I)
```

```
150     D(I) = C(I) **2
160     E(I)= D(I)*F(I)
170     T1= T1 + F(I)
180     T2 = T2 + G (I)
190     T3 = 73 + E(I)
200     PRINT USING 330, E(I), D(I), G(I), C(I), F(I), B(I), A(I), I
210     NEXT I
220     PRINT USING 340 230         PRINT USING 350, T3, T2, T1
240     V= (T3-T2 **2/T1)/(T1-1)
250     R= SQR (V)
260     PRINT
270     PRINT
280     PRINT, V;` = التباين  VARIANCE
285     PRINT
290     PRINT, R;`=
300
310     س2كر    سر2       سركر    سر    كر   الفئة    رقم الفئة
320     ____________________________________________________
330     *****          ****       ****   **   **  **.*...**.*   **
340     _______      ______          _______
350     ******       ****             ***
360     PRINT
370     PRINT
380     DATA 13, 12.5, 15.2, 2, 15.5, 18.5, 3, 18.5, 21.5, 4, 21.5,
390     24.5, 5, 24.5, 27.5, 14, 27.5, 30.5, 21, 30.5, 33.5, 59, 33.5,
400     36.5, 24, 36.5, 39.5, 7, 39.5, 42.5, 6, 42.5, 45.5, 2, 45.5,
410     48.5, 1, 48.5, 51.5, 2
420     END
```

المخرجات

رقم الفئة	الفئات	F $ك_ر$	$\overline{X}$ $س_ر$	$F\overline{X}$ $س_ر ك_ر$	$\overline{X}^2$ $س_ر^2$	$س_ر^2 ك_ر$
1	15.5 – 12.5	2	14	28	196	392
2	18.5 – 15.5	3	17	51	289	867
3	21.5 – 18.5	4	20	80	400	1600
4	24.5 – 21.5	5	23	115	429	2645
5	27.5 – 24.5	14	26	364	676	9464
6	30.5 – 27.5	21	29	609	841	17661
7	33.5 – 30.5	59	32	1888	1024	60416
8	36.5 – 33.5	24	35	840	1225	29400
9	39.5 – 36.5	7	38	266	1444	10108
10	42.5 – 39.5	6	41	246	1681	10086
11	45.5 – 42.5	2	44	88	1936	3872
12	48.5 – 45.5	1	47	47	2209	2209
13	51.5 – 48.5	2	50	100	2500	5000
	المجموع	**150**		**4722**		**$\sum$153720**

التباين = 34.03691

الانحراف المعياري = $\sqrt{34.03691}$

= 5.834116

1/6/5 الانحراف المعياري والمقارنات:

الانحراف المعياري، وكذلك الوسط الحسابي، يتأثران بوحدات القياس، لذلك لا بد من اللجوء إلى مقياس آخر يخلو من وحدات القياس، وهذا ما توصل إليه كارل بيرسون (1857 ,193)، عندما أثت أن نسبة الانحراف المعياري إلى الوسط الحسابي لنفس المجموعة تمثل مقياساً أفضل لمقارنة التشتت بين المجموعتين، ولقد سمي هذا المقياس بمعامل الاختلاف معامل.

معامل الاختلاف Coefficient of Variation

$$\therefore \text{معامل الاختلاف} = \frac{\text{الانحراف المعياري}}{\text{الوسط الحسابي}} \times 100 = \frac{ع}{\overline{س}} \times 100$$

ومعامل الاختلاف نسبة مئوية تخلو خُلَوّاً تاماً من وحدات القياس، ويمكـن اسـتخدامه لمقارنـة التشـتت بـين أي مجموعتين، سواءً بنفس وحدة القياس أم بغيرها.

مثــال:

الوسط الحسابي لمجموعة ما يساوي (150)، بينما كان الانحراف المعياري لنفس المجموعة وبنفس وحدة القياس يساوي (100)، أما الوسط الحسابي لمجموعة ثانية وبوحدة قيـاس مختلفـة فيسـاوي (60)، وانحرافهـا المعيـاري يسـاوي (5)، فأي المجموعتين أكثر تشتتاً؟

الحل:

الانحراف المعياري للمجموعة الأولى يساوي 20 مرة للانحراف المعياري الخـاص بالمجموعـة الثانيـة، إلا أن ذلـك ليس دليلاً على أن المجموعة الأولى هي الأكثر تشتتًأ إذ لا بد من استخدام معامل الاختلاف.

معامل الاختلاف للمجموعة الأولى = $\frac{\text{الانحراف المعياري } 100}{\text{الوسط الحسابي } 150} \times 100 = 6.667\%$

معامل الاختلاف للمجموعة الثانية = $\frac{5}{60} \times 100 = 8.33\%$

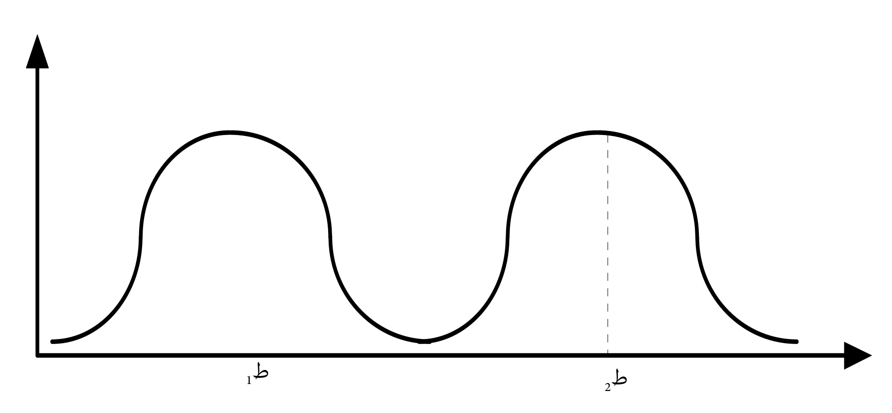
إذاً فالمجموعة الثانية هي الأكثر تشتتاً.

القيم المعيارية Standardized Value

إذا كانت (ظ) هي الوسط الحسابي للمجتمع الذي سحبت منه العينة وإذا كانـت (م) هـي الانحـراف المعيـاري لبيانات المجتمعت.

فهذا يعني أن ($\overline{\text{س}}$) و (ع) هما مقدران لقيمة ظ وم.

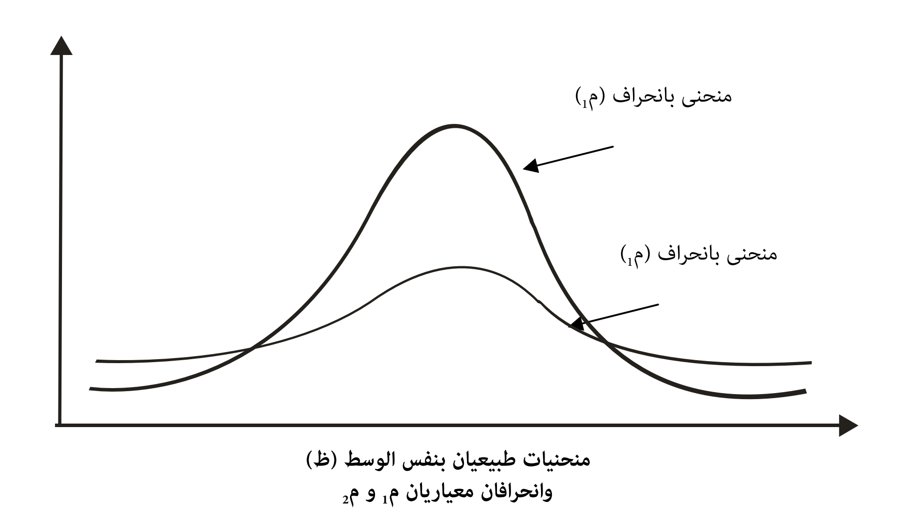
أما ظ و م) فتسميان مَعْلَمي التوزيع الطبيعي، إذا كان التوزيع طبيعياً. وعنـدها يكـون التوزيـع مـتماثلاً حـول الخط العمودي الساقط على ظ، فإذا تحركت ظ إلى اليمين أو اليسار انتقل معها التوزيع، لأنها تمثل المرتكز بالنسبة له.



تحريك الوسط الحسابي -المركز - التوزيع الطبيعي

أما المعلم الثاني (م) فهو مقياس التشتت، لذلك فهي التي تحدد شكل المنحنى بعد أن يكـون موقعـه قـد حـدّد بواسطة (ظ).

إذ كلما قلت (م) زاد ارتفاع المنحنى وقل تشتته، والعكس صحيح.



منحنيات طبيعيان بنفس الوسط (ظ)
وانحرافان معياريان م$_1$ و م$_2$

حيث م$_2$ > م$_1$ زيادة التشتت حول المركز كلما زادت م.

$$ي = \frac{س_ر - \overline{س}}{ع}$$

والقيمة المعيارية هي التي تستخدم للمقارنة بين القيم التي تتبع لمجتمعات مختلفة على أساس عدد الوحدات المعيارية الناتجة بعد التحويل، والتي توضح الترتيب الخاص بكل متغير في مجتمعه اعتماداً على الشكل (4).

مثال:

حصل أحد الطلاب على (81) درجة في أحد الامتحانات التي كان الوسط العام فيها لجميع الطلاب الممتحنين (70) درجة بانحراف معياري قدره (10)، بينما حصل طالب آخر في مؤسسة تعليمية أخرى على (90) درجة في نفس المادة، وكان الوسط الحسابي والانحراف المعياري للمتحنين في المدرسة الثانية هما (75) و (16) على التوالي.

فأي الطالبين أفضل مقارنة بالمجموعة التي يدرس معها ؟

الحل:

ي$_1$ = $\frac{81 - 70}{10} = 1.1$ انحراف معياري

ي$_2$ = $\frac{90 - 75}{16} = \frac{15}{16} = 0.94$ انحراف معياري

إذا فمستوى الطالب الأول هو الأفضل، لأنه يبعد عن الوسط بـ 101 انحراف معياري، بينما يزيد الثاني على الوسط بمقدار 0.94 انحراف معياري.

البرنامج التالي يقوم بحساب القيمة المعيارية لأي رقم من مجموعة أرقام باستخدام المعادلة:

$$Y = \frac{X - M}{R}$$

حيث:

القيمة المعيارية = Y

التغير = X

الوسط الحسابي = M

الانحراف المعياري = R

برنامج لحساب القيمة المعيارية لمجموعة مفردات

```
10    REM
20    DIM  X  (17)
30    S = 0
40    READ  N  REM   عدد القيم
50    PRINT USING 330
60    PRINT USING 340
70    FOR I = 1  TO  N
80    READ  X(I)
90    PRINT USING 350, X(I)
100   S = S +  X(I)
110   NEXT  I
120   PRINT  USING  360
130   PRINT USING  370,  S
140   M = S/N
150   PRINT
160   FOR I – 1 TO N
170   D I X (I) – M
180   T = T + D ** 2
190   NEXT I
200   PRINT
210   PRINT
220   V = T / (N-1)
230   R = SQR (V)
240   PRINT, M;` = Mean الوسط الحسابي
250   PRINT
260   PRINT, R; ` = STANDARD DEVIATION الانحراف المعياري
270   PRINT
280   Y = (X(5) – M)/R
290   PRINT USING 380, Y, X (X)
300   PRINT
310   PRINT
330   :                                            القيم
340   :                                            _____
350   :                                            **
360   :                                            ____
370   :                                            ***  المجموع  TOTAL
380   :            القيمة المعيارية للرقم  **  هي ***.**
390   : DATA 17, 16,22,21,20,23,21,19,15,13,
                23,17,20,29,18,22,16,25
400   END
```

المخرجات:

	القيم
	16
	22
	21
	20
	23
	21
	19
	15
	13
	23
	20
	29
	18
	22
	16
	25
المجموع	340

الوسط الحسابي = 20

الانحراف المعياري = 3.984344

القيم المعيارية للرقم 23 هي 0.753

7.5 العــزوم **Moments**

إذا كانت لدينا مجموعة من القيم تحوي N قيمة، وذلك كالتالي:

$X_1, X_2, X_3, \ldots , XN$

فإن العزم الأول = $\frac{\sum X_1}{N}$

والعزم الثاني = $\frac{\sum X_2}{N}$

والعزم الثالث = $\frac{\sum X_3}{N}$

والعزم الرابع = $\frac{\sum X_4}{N}$

وهكذا يمكن إيجاد أي عزم فمثلاً لو كانت لدينا القيم:

4, 6, 10, 20

فإذا رمزنا للعزم الأول بـ M1، والعزم الثاني بـ M2، والعزم الثالث M3، والعزم الرابع بـ M4...

يمكننا تطبيق القوانين السابقة، وذلك على النحو التالي:

$$M_1 = \frac{\sum X}{N} = \frac{4+6+10+20}{4} = 10$$

$$= \frac{(4)^2 - (6)^2 + (10)^2 + (20)^2}{4} = \frac{552}{4} = 138$$

أما العزم الثاني

$$M_2 = \frac{\sum X^2}{N}$$

أما العزم الثالث فهو

$$M_3 = \frac{(4)^2 - (6)^2 + (10)^2 + (20)^2}{4}$$

$$= \frac{256 + 129 + 10000 + 160000}{4} = 42888$$

من المثال السابق يلاحظ بأن:

- العزم الأول يساوي = المتوسط الحسابي $\overline{X}$ لمجموعة القيم.
- والتباين يساوي الزم الثاني – ناقصاً مربع العزم الأول.

وبتطبيق المثال السابق، يلاحظ بأن العزم الثاني يساوي 138 والعزم الأول يساوي 10.

وبالتالي فإن التباين يساوي 38، $(138-(10)^2)$.

أما إذا كانت البيانات التي لدينا مبوبة في فئات، فإن:

العزم الأول يكون $M_1 = \frac{\sum XF}{N}$

والعزم الثاني $M_2 = \frac{\sum X^2F}{N}$

والعزم الثالث $M_2 = \frac{\sum X^3F}{N}$

والعزم الرابع $M_2 = \frac{\sum X^4F}{N}$

حيث X هنا تمثل مراكز الفئات و F تشير إلى التكرارات وNتشير إلى مجموع التكرار.

ويمكن حساب العزم الأول والثاني والثالث والرابع للبيانات المبوبة في فئات، كما يلي:

مركز الفئة X	F	XF	X^2F	X^3F	X^4F
2	1	2	4	8	16
3	4	12	36	108	324
4	3	12	48	192	768
5	2	10	50	250	1250
المجموع	**10**	**36**	**138**	**558**	**2358**

وبالتالي، فإن العزم الأول، يكون $M_1 = \frac{63}{10} = 3.6$

والعزم الثاني يكون $M_2 = \frac{138}{10} = 13.8$

والعزم الثالث يكون $M_3 = \frac{558}{10} = 55.8$

والعزم الرابع، يكون $M_4 = \frac{2358}{10} = 235.8$

والمتوسط الحسابي x، كما قلنا سابقاً يساوي العزم الأول = 3.6، والتباين يساوي 0.84.

(العزم الثاني – مربع العزم الأول) العزوم حول المتوسط الحسابي X.

وتحسب العزوم حول المتوسط الحسابي لمجموعة من المشاهدات أو البيانات، تحوي N مشاهدة، كالآتي:

العزم الأول حول المتوسط الحسابي يكون $M_1 = \frac{\sum (X - \overline{X})}{N}$

والعزم الثاني حول المتوسط الحسابي يكون $M_2 = \frac{\sum (X - \overline{X})^2}{N}$

والعزم الثالث حول المتوسط الحسابي يكون $M_3 = \frac{\sum (X - \overline{X})^3}{N}$

والعزم الرابع حول المتوسط الحسابي يكون $M_4 = \frac{\sum (X - \overline{X})^4}{N}$

فإذا كانت لدينا القيم التالية: 20 ,10 ,6 ,4

فإنه يمكننا إيجاد العـزوم حـول المتوسـط الحسـابي لهـذه المشـاهدات، فـأولاً، نجـد المتوسـط الحسـابي X لهـذه المشاهدات:

$$X = \frac{4+6+10+20}{4} = \frac{40}{4} = 10$$

ثم يمكن وضعه في جدول ليحسب العزوم حول المتوسط الحسابي

x	$(X-\overline{X})$	$(X-\overline{X})^2$	$(X-\overline{X})^3$	$(X-\overline{X})^4$
4	-6	36	-16	1296
6	-4	10	-64	256
10	-	-	-	-
20	10	100	1000	10000
المجموع	**0**	**152**	**1000** **-280** **720**	**11552**

ومن الجدول السابق، يلاحظ أن العزوم، حول المتوسط الحسابي يمكن حسابها على النحو التالي:

فالعزم الأول حول المتوسط الحسابي، يكون: $M_1 = \frac{0}{4} = 0$

فالعزم الثاني حول المتوسط الحسابي، يكون: $M_2 = \frac{152}{4} = 38$

فالعزم الثالث حول المتوسط الحسابي، يكون: $M_3 = \frac{720}{4} = 180$

فالعزم الثالث حول المتوسط الحسابي، يكون: $M_4 = \frac{1152}{4} = 2888$

ويلاحظ أن هذه الحالة، أن العزم الأول حول المتوسط الحسابي يساوي صفر دائماً.

والعزم الثاني حول المتوسط الحسابي يساوي التباين:

أما إذا كانت بيانات مبوبة في جداول تكرارية، فإن العزوم حول المتوسط الحسابي يمكن حسابها على النحو التالي:

- العزم الأول حول المتوسط الحسابي، يكون $M_1 = \frac{\sum \left(X - \overline{X}\right) F}{N}$
- أما العزم الثاني حول المتوسط الحسابي، يكون $M_2 = \frac{\sum \left(X - \overline{X}\right)^2 F}{N}$
- أما العزم الثالث حول المتوسط الحسابي، يكون $M_3 = \frac{\sum \left(X - \overline{X}\right)^3 F}{N}$
- أما العزم الرابع حول المتوسط الحسابي، يكون $M_4 = \frac{\sum \left(X - \overline{X}\right)^4 F}{N}$

حيث X تشير إلى مركز الفئة و $\overline{X}$ تشير إلى المتوسط الحسابي و F تشير إلى التكرار و N تساوي مجموع التكرار وإذا أخذنا، جدول تكراري فإنه يمكننا حساب العزم الأول حول المتوسط الحسابي، والعزم الثاني حول المتوسط الحسابي، وعلى هذا النمط، يمكن حساب باقي العزوم، وذلك كالآتي:

X	F	XF	$\left(X - \overline{X}\right) F$	$\left(X - \overline{X}\right)^2 F$	$\left(X - \overline{X}\right)^3 F$	$\left(X - \overline{X}\right)^4 F$
1	6	6	6	6	6	6
2	9	18	-	-	-	-
3	4	12	4	4	4	4
4	1	4	2	4	8	16
	20	**40**	**0**	**14**	**12** **-6** **6**	**26**

وبالتالي فإنه يمكننا حساب العزوم حول المتوسط الحسابي على النحو التالي:

- العزم الأول حول المتوسط الحسابي، يكون $M_1 = \frac{0}{20} = 0$
- أما العزم الثاني حول المتوسط الحسابي، يكون $M_2 = \frac{14}{20} = 0.7$

وهذا يساوي التباين

- أما العزم الثالث حول المتوسط الحسابي، يكون $M_3 = \frac{6}{20} = 0.3$

ويمكن حساب باقي العزوم على النحو السابق وحسب القوانين التي ذكرت، وتستخدم العزوم في إيجاد المعامل العزمي للالتواء، وكذلك معامل التفرطح.

8/5 الالتـــواء: Skew ness

يعبر الالتواء عن درجة تماثل توزيع البيانات حول المتوسط الحسابي، أو درجة عدم تماثل المنحنيات التكرارية في حالة البيانات المبوبة في جداول تكرارية، وهناك ثلاثة مقاييس للالتواء، هي:

- المعامل الأول للالتواء = (المتوسط الحسابي – المنوال) / (الانحراف المعياري)

- المعامل الثاني للالتواء = ٣(المتوسط الحسابي – الوسيط) / (الانحراف المعياري)

- المعامل الثالث للالتواء = (العزم الثالث حول المتوسط الحسابي) / (الانحراف المعياري)

والمنحنيات التكرارية بعضها متماثل، وبعضها غي متماثل أو ملتو.

- إذا كان المنحنى متماثل، فإن الوسط الحسابي والمنوال متساويين، ونستخدم هذه الخاصية كأساس لقياس الالتواء، ونأخذ العامل الأول للالتواء.
- الفرق بين المتوسط الحسابي والمنوال يساوي صفر، إذا كان المنحنى متماثلاً، ويكون قريباً من الصفر إذا كان المنحنى قريباً من التماثل.
- ويكون الالتواء كبيراً إذا كان الفرق كبيراً وهذا القرق نقيسه بالنسبة للانحراف المعياري للتوزيع.

- إذا كان المتوسط الحسابي أكبر من المنوال، كان الالتواء، موجباً ويكون المنوال أصغر من المتوسط الحسابي في حالة الالتواء الموجب.
- وفي حالة الالتواء الموجب أي أن المنحنى ملتو إلى اليمين، وفي هذه الحالة تتركز التكرارات عند القيم الصفري فيصعد المنحنى بسرعة ويهبط ببطء.
- أما المنحنى الملتو إلى اليسار، أي التواء سالب، فإن التكرارات تتركز عند القيم الكبرى، فالمنحنى يصعد ببطء ويهبط بسرعة نستخلص.
- إذا كان معامل الالتواء مساوياً الصفر، فإن البيانات تكون تماثلة حول المتوسط الحسابي، أو إن المنحنى التكراري متماثل.
- أما إذا زاد معامل الالتواء عن الصفر، فإن المنحنى التكراري ملتو إلى اليمين.
- أما إذا قلّ معامل الالتواء عن الصفر، فإن المنحنى التكراري ملتو إلى اليسار.

وحساب المعامل الأول للالتواء، والمعامل الثاني للالتواء يتم بسهولة لمعرفتنا لكيفية إيجاد المتوسط الحسابي، والمنوال والوسيط، وكذلك الانحراف المعياري.

مثـــال:

المطلوب حساب المعامل العزمي للالتواء في المثال السابق، حيث وجدنا أن العزم الثالث حول المتوسط الحسابي = 3.

والعزم الثاني حول المتوسط الحسابي = التباين = 7

ومن هنا يمكن إيجاد الانحراف المعياري:

$$\sigma = \sqrt{\text{التباين}} = \text{التباين}$$

$$= \sqrt{0.7}$$

$$= 0.83$$

$$\text{المعامل العزمي للالتواء} = \frac{\text{العزم الثالث حول المتوسط الحساب}}{(\text{الانحراف المعياري})^3}$$

العامل العزمي للالتواء =

$$= \frac{m_3}{(\sigma\sigma)}$$

$= .3$

$= (.83)^3$

$= 0.52$

أي أن المنحنى التكراري يكون ملتو إلى اليمين.

1/8/5 تطبيقات الالتواء على الحاسوب

الالتواء Skewness

الالتواء: هو عدم تماثل التوزيع بالنسبة لأي خط عمودي، لذلك يتم قياس الالتواء لتحديد ما إذا كان هناك التواء موجب أو سالب ودرجة الالتواء لتوضيح شكل منحنى توزيع البيانات، ويستخدم في ذلك العزم الثالث.

إذا تساوى مجموع مكعبات الانحرافات الخاصة بالقيم التي تقل عن الوسط الحسابي بمجموع مكعبات انحرافات القيم التي تزيد على الوسط الحسابي، يكون العزم الثالث صفراً. إذاً فمعامل الالتواء للتوزيع التماثل يساوي صفراً، طالما أنه يعتمد على العزم الثالث.

$$ل_3 = \frac{\sum (س_ر - \bar{س})^3}{ن}$$

لقد استخدمت القيمة المعيارية للالتواء، وسميت بمعامل الالتواء Coefficient of Skewness للتخلص من وحدات القياس، فقد عرف معامل الالتواء.

$$\sqrt{ت_1} = \frac{ل3}{ع3}$$

ويلاحظ أن المقام هو مكعب الانحراف المعياري حتى يخلو معامل الالتواء خلوّاً تاماً من وحدات القياس أو مقياس الرسم، كذلك يمكن تعريف معامل الالتواء بأنه:

$$ت_1 = \frac{(ل3)^2}{(ع3)^2}$$

وبما أن ع = $\sqrt{ل_2}$ فإن:

$(ع_3)^2 = ل_2^3$ وبذلك تكون: $ت_1 = \frac{ل_3^2}{ل_2^3} =$ $\frac{\text{مربع العزم الثالث}}{\text{مكعب العزم الثاني}}$

كما أن:

$$\sqrt{ت_1} = \frac{ل_3}{ل_2\sqrt{ل_2}}$$

مثال: أوجد معامل الالتواء للتوزيع التالي:

الفئات	$ك_ر$	$س_ر$	$س_رك_ر$	$س_ر - \bar{س}$	$(س_ر - \bar{س})^2$	$(س_ر - \bar{س})^2 ك_ر$	$(س_ر - \bar{س})^3 ك_ر$
7 – 3	1	5	5	12.24-	149.8176	149.8176	1833.7674-
11 – 7	3	9	27	8.24-	67.8976	203.6928	1678.4286-
15-11	11	13	143	4.24-	17.9776	197.7536	838.47526-
19-15	20	17	340	0.4-	0.0576	1.152	0.27648-
23-19	9	21	189	3.76	14.1376	127.2384	478.41638
27-23	4	25	100	7.76	60.2176	240.8704	1869.1542
31-27	2	29	58	11.76	138.2976	276.5952	3252.7596
المجموع	**50**				**448.4032**	**1197.120**	**1249.3824**

$\bar{س} = 17.24$

$$ل_2 = \frac{\sum (س_ر - \bar{س})^2 ك_ر}{ن} = \frac{1197.120}{50} = 23.942$$

$$ل_3 = \frac{\sum (س_ر - \bar{س})^3 ك_ر}{ن} = \frac{1249.3829}{50} = 24.988$$

$$\sqrt{ت_1} = \frac{ل_3}{ل_2\sqrt{ل_2}} = \frac{24.988}{23.942\sqrt{23.942}}$$

$\therefore \sqrt{ت_1} = 0.213$

$ت_1 = 0.0455$

البرنامج التالي يحسب معامل الالتواء والذي سبق الحديث عنه مستخدماً في ذلك التوزيع التكراري السابق باستخدام المعادلة:

$$W = Q^2$$

$$Q = \frac{L_3}{L_2\sqrt{L_2}}$$

العزم الثاني = L2

العزم الثالث = L3

برنامج لحساب الالتواء

```
10	REM
20	DIM  A(7), B(7), C(7), D(7), E(7), F(7), G(7), H(7), K(7), L(7)
30	P = 0
40	V = 0
50	R = 0
60	S = 0
70	T = 0
75	D1 = 0
80	READ  N  REM عدد المشاهدات
90	FOR I  = 1 TO N
100	READ A (I), B (I), F(I), REM   الحد الأدنى، الحد الأعلى، التكرار
110	T = T + F (I)
120	C(I) = (A(I) + B(I))/2
130	D(I) = C(I) * F(I)
140	D1 = D1 + D(I)
```

```
150        NEXT I
160        M = D1/T
170        FOR I = 1 TO N
180        E (I) = C(I)-M
190        K (I) = (E(I)) **2
200        L (I) = (C(I) – m**2* F(I)
210        G (I) = (C(I)-M)**3*F(I)
220        H(I) = C(I) * F(I)
230        P = P + F (I)
240        V = V+H (I)
250        R = R + K (I)
260        S= S + L (I)
270        V = V +G(I)
280        NEXT I
285        PRINT USING  305
290        PRINT USING 300
295        PRINT USING 305
300        الفئة  كر  سر سر كر (سر-س-) (سر-س-)2 (سر-س-)3 كر
305        : ______________________________________________
310        PRINT
320        FOR I = 1 TO N
330        PRINT USING 340, G(I), L(I), K(I), KI), E(I), H(I), C(I),
           F(I), F(I), B(I), A(I)
340        :******.***  ****.***  ***.***   ***.***  ***  *** ** **.*=**=
350        NEXT I
360        PRINT USING 370
370        : _____    ______    ____    ___    ____
380        PRINT
390        PRINT USING 400, V, S, R, V, P
400        : ****.****  ****.***   ****.***   ****   ***  المجموع
401        L2 = S/P
402        L3 – V/P
403        Q = L3/ (L2*SQR (L2))
405        W = Q **2
420        PRINT USING 430, W
430        : COEFFICIENT OF **.***** SKEWNESS عامل الالتواء
440        PRINT
450        PRINT
460        DATA  7,3,7,1,7,11,3,11,14,15,,19,20,
                      19, 23, 9, 23, 9, 23, 27, 4, 27, 31, 2
480        END
```

المخرجات

الفئات	ك ر	س ر	س ر ك ر	س ر -$\bar{س}$	$(س_ر - \bar{س})^2$	$(س_ر - \bar{س})^2$ ك ر	$(س_ر - \bar{س})^3$ ك ر
7.0-3.0	1	5	5	12.24-	149.8176	149.8176	1833.7674-
11.0-7.0	3	9	27	8.24-	67.8976	203.6928	1678.4286-
15.0-11.0	11	13	143	4.24-	17.9776	197.7536	838.47526-
19.0-15.0	20	17	340	0.4-	0.0576	1.152	0.27648-
23.0-19.0	9	21	189	3.76	14.1376	127.2384	478.41638
27.0-23.0	4	25	100	7.76	60.2176	240.8704	1869.1542
31.0-27.0	2	29	58	11.76	138.2976	276.5952	3252.7596
المجموع	**50**		**862**		**448.4032**	**1197.120**	**1249.3824**

معامل الالتواء = 0.0455

أما في حالة التوزيعات ذات الفئات المفتوحة، فيمكن استخدام الربيعات والمدى الربيعي لتقدير الالتواء.

فالفرق بين الوسيط والربيع الأعلى يساوي الفرق بين الوسيط والربيع الأدنى في حالة التماثل، أي أن:

,75% - ,25 % = ,50% - ,150% في حالة التماثل.

وأما إذا كان الالتواء سالباً، فيكون الوسيط أقرب للربيع الأعلى، والعكس صحيح لذلك يعـرف معامـل الالتـواء في هذه الحالة، بأنه:

$$ت_1 = \frac{(\%25, - \%50,) - (\%50, - \%75,)}{\%25, - \%75,}$$

$$= \frac{\%25, + \%5092 - \%75,}{\%25, - \%75,}$$

مثـــال:

استخدم البيانات التالية لتقدير الالتواء بطريقة الربيعات:

الفئات	ك ر F	المتجمع التكراري الصاعد
7-3	1	1
11-7	3	4
15-11	11	15
19-15	20	35
23-19	9	44
29-23	4	48
31-27	2	50
	50	

$$\text{ون\%} = ح_1 + \frac{\left(\frac{\text{و ن}}{100} - ن\right) ط}{ك}$$

$$17 = \frac{4\times\left(15 - 25\right)}{20} = 15 = \%50,$$

$$14.09 = \frac{4\times\left(4 - 12.5\right)}{11} = 11 = \%25,$$

$$20.11 = \frac{4\times\left(35 - 37.5\right)}{9} = 19 = \%75,$$

$$0.033 = \frac{14.09 + 17\times 2 - 20.11}{14.09 - 20.11} = ت_1$$

البرنامج التالي يقوم بحساب الآتي:

- الربيع الأعلى.
- الربيع الأدنى.
- الوسيط.
- الانحراف الربيعي.
- معامل الالتواء.

وهو يقوم بحساب معامل الالتواء باستخدام طريقة الربيعات باستخدام المعادلة:

$$Q = \frac{(Q_1 - 2Q_2 + Q_3)}{Q_1 - Q_3}$$

حيث:

معامل الالتواء = Q

الربيع الأعلى = Q1

الوسيط = Q2

الربيع الأدنى = Q3

برنامج لحساب الالتواء Skewnes بطريقة الربيعات

```
10    REM
20    DIM A(7), B(7), C(7), D(7), E(7), F(7), G(7)
30    T = 0
40    READ    N
50    FOR I = 1  TO N
60    READ A (I), B(I), F(I) REM الحد الأعلى، الحد الأدنى، التكرار
70    T =  T  +  F(I)
80    NEXT  I
90    C (1) =   F (1)
100   FOR I = 2  TO N
110   C (I) = C (I-I)+ F (I)
120   NEXT   I
130          رقم الفئة    الفئة    التكرار    المتجمع التكراري الصاعد
140   :      ________________________________________
150   :      ***        **      **.*-**.***        **
160   :      ________________________________________
170   :                  ***    summation    المجموع
```

```
180      :
190      PRINT USING 150
195      PRINT USING 130
200      PRINT USING 140
210      PRINT USING 150
220      PRINT
230      FOR   I =  1  TO   N
240      PRINT USING 160, C(I) , F(I), B(I), A(I); I
250      NEXT I
260      PRINT USING 170
270      PRINT USING 180, T
280      PRINT
290      S1 = T * 3/4
295      S2 = T/2
300      S3 = T/4
310      FOR I = 1  TO  N
320      IF C(I) > S1  THEN 340
330      J = I
340      NEXT   I
342      FOR I = 1  TO  N
344      IF C(I) > S2 THEN  348
346      K = I
348      NEXT  I
350      FOR  I  =1  TO N
360      IF C (I) > S3 THEN 380
370      V = I
380      NEXT  I
400      PRINT
450      GO  SUB  610
460      Q1 FNA  (L1, S1, O1, P1, W1)
470      PRINT, Q1' `=  الربيع الأعلى  ,  QUARTILE
480      PRINT
485      GO SUB 651
490      Q2 = FNA (L2, S2, Q2, P2, W2)
495      PRINT, Q2;` =  الوسيط  MEADIAN
500      PRINT
510      GO  SUB  660
```

```
320     Q3 = FNA
530     PRINT
540     PRINT
550     PRINT
560     Y = (Q1-Q3)/2
570     PRINT , Y;` =   الانحراف المعياري
572     PRINT
573     PRINT
575     Q = (Q1 – 2*Q2 + Q3) / (Q1-Q3)
580     :                              معامل الالتواء = ****.*
585     PRINT
590     PRINT
600     STOP
610     L1 = A (J+1)
620     O1 = C (J)
630     P1 = B (J) – A(J)
640     W1 = F (J+1)
650     RETURN
651     L2 = A (K+1)
652     O2 = C (K)
653     P2 = B (K) – A(K)
654     W2 = F (K+1)
655     W2 = F (K+1)
655     PETURN
660     L3 = A (V+1)
670     O3 = C(V)
680     P3 = B (V) – A (V)
690     W3 = F (V+1)
700     RETURN
710     DEF FNA (L, X, O, P, W) = L + (X-O)*P)/W
720     DATA 7,3,7,1,7,11,3,11,15,11,15,19,20,19,
730     DATA 23,9,23,27,4,27,31,2
750     END
```

المخرجات

رقم الفئة	الفئات	التكرار	المتجمع التكراري الصاعد
1	7.0 – 3.0	1	1
2	11.0 – 7.0	3	4
3	15.0 – 11.0	11	15
4	19.0 – 15.0	20	35
5	23.0 – 19.0	9	44
6	27.0 – 23.0	4	48
7	31.0 – 27.0	2	50
المجموع		**50**	

الربيع الأعلى = 20.1111
الوسيط = 17
الربيع الأدنى = 14.09091
الانحراف الربيعي = 3.010095
معامل الالتواء = .0336

9/5 التفرطح وتطبيقاته على الحاسوب

التفرطح: Kurtosis

التفرطح عكس التدبب، فالتوزيع المتفرطح هو الذي يكون أقل ارتفاعاً من التوزيع الطبيعي.

لذلك فالتوزيع المتفرطح يتميز بمعامل اختلاف أكبر من الطبيعي.

ويعرف معامل التفرطح بأنه نسبة العزم الرابع إلى مربع العزم الثاني، أي أن

$$ت_2 = \text{معامل التفرطح} = \frac{ل_4}{(ل_2)^2}$$

ومعلوم أن العزم الثاني هو التباين Variance لذلك يزداد المقام كلما ازداد التشتت.

يلاحظ أن بسط معامل التفرطح وكذلك مقامه لا يكونان سالبين، لذلك تكون قيمة التفرطح موجبة في جميع الحالات، ولا تساوي صفراً، إلاّ إذا كانت جميع القيم متساوية.

مثـال: أوجد التفرطح للبيانات أدناه:

الفئات	$ك_ر$	$س_ر - \bar{س}$	$(س_ر - \bar{س})^2 ك_ر$	$(س_ر - \bar{س})^4 ر$
7 – 3	1	12.24-	149.8176	22445.31327
11 – 7	3	8.24-	203.6928	13830.25226
15 – 11	11	4.24-	197.7536	3555.135119
19 – 15	20	0.24-	1.152	0.0663552
23 – 19	9	3.76	127.2384	1798.864
27 – 23	4	7.76	240.8704	14504.6374
31 - 27	2	11.76	276.5952	38252.45233
المجموع	**50**		**1197.120**	**94386.69**

$\bar{س} = 17.24$

$$ل_2 = \frac{1197.120}{50} = 23.942$$

$$ل_4 = \frac{94386.69}{50} = 1887.734$$

$$ت_2 = \text{معامل التفرطح} = \frac{ل_4}{(ل_2)^2} = 3.2931$$

التفرطح يعبر عن درجة تدبب المنحنى التكراري، وتقاس درجة التدبب على النحو التالي:

$$\text{المعامل العزمي للتفرطح} = \frac{\text{العزم الرابع حول المتوسط الحسابي}}{\text{(الانحراف المعياري)}^4}$$

وبمعنى آخر

المعامل العزمي للتفرطح = $\frac{m\,4}{(\sigma)4}$

- إذا كان المعامل العزمي للتفرطح يساوي 3، فإن المنحنى يكون معتدل التدبب.
- أما إذا زاد المعامل العزمي للتفرطح أقل من 3، فإن هذا يدل على أن المنحنى التكراري مفرطح.
- ويمكن حساب العزم الرابع حول المتوسط الحسابي، كما هو مبين في الجدول السابق، وباستعمال ما في الجدول نجد أن:

$$\frac{\sum (x - \overline{x})^4 \, F\,26}{N}$$

إذن العزم الرابع المتوسط الحسابي يكون:

$$= \frac{26}{20}$$

وبما أن العزم الثاني حول المتوسط الحسابي = التباين = 0.7. ويمكن إيجاد الانحراف المعياري، كالآتي:

(σ) (التباين)^2 $= (0.7)^2$

المعامل العزمي للتفرطح = $2.65 = \frac{1.36}{0.49}$

ونظراً لأن المعامل يقل عن 3، فإن هذا يدل على أن المنحنى التكراري لهذا التوزيع يميل إلى التفرطح.

البرنامج التالي يقوم بحساب الآتي:

- العزم الثاني. - العزم الثالث.

- العزم الرابع. - التفرطح.

- الالتواء.   - الوسط الحسابي

- معامل التفرطح.

وباستخدام المعادلات

$$M_2 = \frac{P}{T}; \quad M_4 = \frac{R}{T}; \quad K_1 = \frac{M_4}{(M2)_2} \quad S_1 = \frac{M_4}{(M2)_3}$$

حيث:

العزم الثاني = M2

العزم الرابع = M4

التفرطح = K1

الالتواء = S1

مجموع مربعات الانحرافات مرجحة بالتكرارات = P

مجموع التكرارات = T

برنامج لحساب التفرطح

```
10      REM
20      DIM  A(7), B(7), C(7), D(7), E(7), F(7), G(7), H(7), K(7), L(7)
22      P = 0
23      Q = 0
24      R = 0
25      S = 0
30      T = 0
35      D1 = 0
40      READ  N
50      FOR I  = 1 TO N
60      READ A (I), B(I), F(I)
70      T = T + F(I)
80      C(I) = (A(I) + B (I))/2
90      D (I) = C(I) * F(I)
100     D1 = D1 + D (I)
110     NEXT   I
```

```
120      M = D1/T
130      FOR  I  = 1  TO  N
140      E (I) = C(I) – M
150      L (I) = (C(I) – M**2 *F (I)
170      H (I) = (C(I) – M** 4*F(I)
180      P = P +L (I)
190      R = R + H (I)
200      NEXT  I
210      PRINT USING 250
220      PRINT USING 240
230      PRINT USING 250
240      :     كر(سر-س-)4 كر(سر-س-)2 (سر-س-) كر(سر-س-) الفئة
250      :  ____________________________________________
260      : FOR I = 1 TO  N
270      PRINT
280      PRINT USING 290, H (I), L (I) , E(I), F(I), B (I), A(I)
290      :  ******.***     ****.***    ***.***   ***    ***_***
300      NEXT I
310      PRINT USING 320
320      :  _______       ___________    ___________
330      PRINT
340      PRINT USING 350,  R,  P, T
350      : *******.***        *****.***          المجموع***
360      M2 = P/T     REM SECOND MOMENT
370      M4 = R/T     REM FOURTH   MOMENT
380      K1 = M4/M2**2   REMKURTOSIS
400      S1 = M4/M2 ** 3  REM  SKEWNESS
420      PRINT, M2;` =  العزم الثاني
430      PRINT, Q1;` = العزم الثالث
440      PRINT, M4;` = العزم الرابع
450      PRINT , K1; ` =  التفرطح
460      PRINT, S1;` =  الالتواء
470      PRINT
480      PRINT
490      PRINT, M; ` = الوسط الحسابي
500      PRINT
```

```
510        PRINT, K1;` =  معامل التفرطح
520        PRINT
530        DATA 7,3,7,1,7,11,3,11,15,11,15,19,20,
540        DATA 19,23,9,23,27,4,27,31,2
550        END
```

المخرجات

الفئات	ك ر	س ر -س̄	(س ر -س̄)² ك ر	(س ر -س̄)⁴ ر
7 – 3	1	-12.240	149.817	22445.240
11 – 7	3	-8.240	203.692	13830.180
15 – 11	11	-4.240	197.753	3555.097
19 – 15	20	-0.240	1.152	0.066
23 – 19	9	3.760	127.239	1798.864
27 – 23	4	7.760	240.871	14504.700
31 - 27	2	11.760	276.595	38252.570
المجموع	**50**		**1197.119**	**94386.690**

العزم الثاني = 23.94238

العزم الثالث = 0

العزم الرابع = 1887.734

التفرطح = 3.293108

الالتواء =0.137543

الوسط الحسابي = 17.23999

معامل التفرطح =3.293108

الوحدة السادسة

الأرقام القياســة

The Index Numbers

إن الأرقـام القياسـية مفيـدة جـداً بالنسـبة لاتخـاذ القـرارات وغعـداد الخطـط الاقتصـادية، والسياسات المختلفة، ولذلك فإن الأرقام القياسة تعتبر أداة إحصـائية تمكننـا مـن المقارنـة بـين قيمتـين تمـثلان كميـات أو أسـعار أو أيـة متغـيرات أخـرى في فـترات زمنيـة مختلفـة، فهـي تسـتخدم لقيـاس الاختلافات أو التغيرات النسبية في قيم، وبالتالي اتجاهات مجموعة من التغيرات المتشـابهة، وكـما قلنـا سابقاً، قد تكون هذه المتغيرات أسعار السلع أو كميات الإنتاج أو كميات الاستيراد أو التصدير...الخ.

والرقم القياسي هو رقم نسبي يقيس التغير بين قيمتين من قيم الظاهرة ونبدأ

1/6 الرقم القياسي النسبي البسيط

وهو عبارة عن معدل التغير المئوي في قيمة ما بالنسبة لقيمة أخرى مأخوذة كأساس.

يمكن الحصول على الرقم النسبي البسيط بقسمة القيمة الأولى على ÷ القيمـة الثانيـة المـأخوذة كأساس، وذلك لقياس التغير.

والرقم النسبي البسيط الذي ينتج، يمكن أن نسميه "الرقم القياسي البسيط".

مثال:

إذا كان سعر سلعة ما في سنة الأساس (1970)مسـاوياً لــ P0 وأن سـعرها في سـنة المقارنـة 1975) كان P1 .

فلكي يمكن لنا إيجاد الرقم القياسي البسيط أو السعر النسبي للسلعة في سنة المقارنة (1975).

الرقم القياسي البسيط (أو السعر النسبي)=

$$\frac{\text{سعر السلعة في سنة المقارنة}}{\text{سعر السلعة في سنة الأساس}} \times 100$$

مثـال:

لو أن سعر أحدى السلع في سنة 1970 كان 40 درهـماً "فلسـاً وأن سـعرها في سـنة 1975 كـان 50 درهماً "فلساً".

المطلوب: إيجاد الرقم القياسي البسيط باعتبار سنة 1975 سنة أسا.

السعر النسبي للسلعة في سنة 1975 أو الرقم القياسي البسيط =

$$= 100 \times \frac{50}{40}$$

$$125 =$$

أي أن سعر السلعة في سنة 1975 (سنة المقارنة) قد ارتفع 25% عـن سـعرها في سـنة 1970 (سـنة الأساس).

أي أن سع سنة الأساس هو السعر الأصلي وهو يمثل 100.

وأن أي زيادة أو نقص مقدارها وحدة واحدة في السعر النسبي لأي سـنة لاحقـة لسـنة الأسـاس تمثل 1% بالزيادة أو بالنقص عن السعر الأصلي.

مثال:

إذا كان الرقم القياسي البسيط لـنفس السـلعة يسـاوي 97 "فـإن هـذا يعنـي أن السـعر في سـنة المقارنة أقل من السعر في سنة الأساس بواقع 3%".

المطلوب: إيجاد الرقم القياسي البسيط للإنتاج في سنة المقارنة؟

الحل:

$$\text{الرقم القياسي البسيط في سنة المقارنة} = \frac{\text{كمية الإنتاج في سنة المقارنة}}{\text{كمية الإنتاج في سنة الأساس}} \times 100$$

مثال:

مصنع يود مقارنة إنتاجه في سنة 1975 وإنتاجه في سنة 1975 وأن كمية الإنتاج كانت 300 طن في سنة 1975، ثم أصبحت (600) طن في سنة 1978.

المطلوب: إيجاد الرقم القياسي والبسيط للإنتاج في سنة 1978؟

الحل:

$$200 = 100 \times \frac{600}{300}$$

وهذا يعني أن الإنتاج قد ازداد بواقع 100% عما كان عليه في سنة 1975.

"ويمكن إيجاد الرقم القياسي البسيط للكميات المستوردة من سلعة أو الكميات المصدرة من سلعة ما، بنفس الطريقة السابقة".

ويعاب على طريقة الرقم القياسي البسيط والمركب، أنها:

- لا تعطي وزن للأهمية النسبية للسلع الداخلة في تركيب الرقم القياسي.
- الأهمية النسبية للسلع تختلف، وبالتالي يجب أن يؤخذ بالاعتبار تلك الأهمية للسلع.
- لذلك فإن هذا الرقم لا يعتبر من الأرقام المفضلة لقياس التغيرات في الأسعار.
- والبديل لهذا أخذ المتوسط الحسابي للأرقام القياسية البسيطة لمختلف السلع.

فإذا أخذت سنة 1970 كسنة أساس، فإنه يمكن حساب مناسيب السعر لكل سلعة على حجه.

الرقم القياسي للجبنة = $110.3 = 100 \times \frac{2.4}{2}$

الرقم القياسي البسيط للزبدة = $110.4 = 100 \times \frac{3.5}{3}$

الرقم القياسي البسيط للزيتون = $110.4 = 100 \times \frac{4}{3.5}$

وبالتالي فإن الرقم القياسي التجميعي البسيط، يكون:

$$\frac{125 + 110.3 + 110.4}{3} = \frac{445.7}{3} = 148.5$$

وبذلك تمكّنا من التغلب على مشكلة الأهمية النسبية للسلع الداخلة في تركيب الرقم القياسي بواسطة حساب الرقم القياس التجميعي المرجح والذي يستخدم أوزان مناسبة من ترجيح سعر كل سلعة "وفي العادة تستخدم الكميات المنتجة أو المشتراه أو المباعة كأوزان لترجيح أسعار السلع الداخلة في تركيب الرقم القياسي عند حساب الرقم القياسي للأسعار.

وعند استخدام للكميات كأوزان ترجيح يمكن استخدام كميات سنة الأساس أو كميات سنة المقارنة.

أما إذا استخدمنا رقم قياسي تجميعي مرجح للكميات، فإنه يمكن استخدام الأسعار لهذه السلع كأوزان لترجيح كميات هذه السلع.

وأيضاً يمكن استخدام إمّ: أسعار سنة الأساس، أو أسعار سنة المقارنة

1/2/9 رقم لاسبير **Laspare**

إذا ما استخدمنا الكميات في (سنة الأساس) لكل سلعة من السلع الداخلة في حساب الرقم القياسي كأوزان لترجيح أسعارها، فإننا نسمي هذا الرقم برقم لاسبير.

ويمكن حسابه باستخدام المعادلة التالية $L_1 = \frac{\sum P1q0}{\sum P0q0} \times 100$

حيث P190 عبارة عن سعر سنة المقارنة مرجحاً بكميات سنة الأساس.

P090 عبارة عن سعر سنة الأساس مرجحاً بكميات سنة الأساس.

ويمكن حساب رقم لاسبير، بالمثال التالي:

السنة	الجبنة		الزبدة سعر الكيلو غرام بالدينار	الكمية المباعة كيلو غرام
	سعر الكيلوغرام بالدينار	الكمية المباعة كيلو غرام		
2000	5	20	3	40
2005	3	40	4	30

وبالتالي يمكن حساب رقم لاسبير، على النحو التالي:

$$L_1 = \frac{\sum P1q0}{\sum P0q0} \times 100$$

$$= \frac{3\times 20 + 3\,4\times 40}{5\times 20 + 3\times 40}\times 100$$

$$= \frac{3\times 20 + 4\times 40}{5\times 20 + 3\times 40}\times 100 \quad = \frac{60+160}{100+120}\times 100$$

$$= 100$$

2/2/6 رقم باشي **Pachi**

أما إذا استخدمنا كميات سنة المقارنة لكل سلعة من السلع الداخلة في حساب الرقم القياسي كأوزان لترجيح أسعارها، فإن الرقم الناتج يسمى برقم باشي.

يحسب رقم باشي بالمعادلة التالية =

$$B_1 = \frac{\sum P1q1}{\sum P0q1} \times 100$$

حيث P091 تعني سعر سنة الأساس مرجحاً بكميات سنة المقارنة.

وP191 تعني سعر سنة المقارنة مرجحاً بكميات سنة المقارنة، ويمكن استخدام المثال الآتي لحساب رقم باشي.

لسنة	الجبنة		الزبدة	
	سعر الكيلوغرام بالدينار	ا الكمية المباعة كيلو غرام	بالدينار سعر الكيلو غ رام	الكمية المباعة (كيلو غرام)
2000	5	20	3	1970
2005	3	40	4	1975

ويمكن حساب رقم باشي على النحو التالي $$B_1 = \frac{\sum P191}{\sum P091} \times 100$$

$$B_1 = \frac{3 \times 40 + 4 \times 30}{5 \times 40 + 3 \times 30} \times 100$$

$$= \frac{120 + 120}{200 + 90} \times 100 \qquad = \frac{240}{290} \times 100 = 82.8$$

ورقم لاسبير Laspar يميل دائماً إلى تكبير قيمة الرقم المقدر.

بينما رقم باشي يميل إلى تصغير الرقم المقدر.

وهذا يعني أن رقم لاسبيرمتحيز إلى أعلى.

بينما رقم باشي متحيز إلى أسفل.

وللتخلص من هذا التحيز، نستخدم رقم آخر، وهو متوسط رقمي "لاسبير وباشي"، وهذا الرقم يسمى برقم مارشال أوجورث، ويمكن حسابه بالمعادلة التالية:

$$M - E_1 = \frac{\sum P_1 \ (q0 + q1)}{\sum P_0 \ (q0 + q1)} \times 100$$

حيث P1، 91 تعني أسعار وكميات سنة المقارنة.

و P0 ، 90 تعني أسعار وكميات سنة الأساس، بالترتيب.

3/2/6 رقم مارشال أدجوورث

وإذا ما حسبنا هذا الرقم من المثال السابق، فإنه يكون كالآتي:

$$M - E_1 = \frac{460}{510} \times 100 = 90.4$$

وبالتالي، فإن رقم لاسبير = 100

ورقم باشي = 82.8

ورقم مارشال أدجوورث = 90.4

لذلك فإن الرقم الأخير (رقم مارشال أدجوورث) أدى إلى التخلص من التحيـز، إلا أن هنـاك رقـماً آخر أمثل يسمّى (رقم فيشر Fisher) الأمثل.

وهذا الرقم عبارة عن المتوسط الهندسي لرقمي لاسبير وباشي (لاسبير وباش).

ويمكن حسابه باستخدام المعادلة التالية: المثال السابق، حيث:

$$FI = 100\sqrt{\frac{\sum P190}{\sum P090} \times \frac{\sum P191}{\sum P091}}$$

ويمكن حسابه باستخدام المثال السابق، حيث:

$$FI = 100\sqrt{\frac{\sum 220}{\sum 220} \times \frac{\sum 240}{\sum 290}}$$

$$= 100\sqrt{\frac{52800}{63800}} = 91$$

3/6 تغيير سنة الأســـــاس

من الممكن تغيير سنة الأساس، بسنة أخرى، وهذا يتم عندما يراد مقارنة رقمين قياسيين.

لكي تتم المقارنة على أساس دقيق، يتم توحيد سنة الأساس للرقمين القياسيين، وذلك بتغيير سـنة الأساس لأحد الرقمين لتطابق سنة الأساس للرقم الثاني.

مثـــــال:

نفرض أنه لدينا سلسلة من الأرقام القياسية لنفقة المعيشة محسوبة على أساس أنه سنة الأساس 1960-100 كالآتي:

الرقم القياسي	السنة
100	1960
101	1961
104	1962
108	1963
110	1964
112	1965

وأردنا مقارنة الأرقام القياسية لنفقة المعيشة لبلد آخر محسـوباً رقمـه القيـاسي لنفقـة المعيشـة على أساس سنة 1963 كسنة أساس.

يتم حساب الرقم لنفقة المعيشة الجديد على أساس 1963 كسنة أساس، وذلك بقسمة الأرقـام السـابقة عـلى 108، وهو الرقم القياسي لسنة 1963 مضروبة × في 100، كالآتي:

الرقم القياسي	السنة
$\frac{100}{108} \times 100 = 92.6$	1960
$\frac{101}{108} \times 100 = 93.6$	1961
$\frac{104}{108} \times 100 = 96.3$	1962
$\frac{108}{108} \times 100 = 100$	1963
$\frac{110}{108} \times 100 = 101.9$	1964
$\frac{112}{108} \times 100 = 103.7$	1965

الوحدة السابعة

الارتبـــــــاط

CORRELATION

الارتباط هو دراسة العلاقة بين متغيرين أو أكثر.

والارتباط يهتم بالتغير المشترك في عدد من الظواهر الـذي قـد يعـود عـلى وجـود علاقـة سـببية بينهما، وبحيث يمكننا أن نميز بوضوح بين السبب والنتيجة.

1/7 الارتباط السببي **Casual Correlation**

يكون الارتباط السببي إذا كان السبب والنتيجة واضحين وفي اتجاه معين. فزيـادة الإنتـاج سـببها زيادة السماد

2/7 الارتباط التبادلي **Mutual correlation**

يكون الارتباط تبادلياً عندما تؤثر كل من الظاهرتين (أو اكثر) في بعضهما البعض، بحيث لا يمكن التمييز بين السبب والنتيجة.

مثلاً زيادة كمية النقود تسبب في ارتفاع الأسـعار، في حـين يـرى آخـرون أن ارتفـاع الأسـعار هـو الذي يؤدي إلى كمية النقود، ويمسى هذا النوع من الارتباط- بالارتباط التبادلي .Mutual Creation

3/7 الارتباط الوهمي **Spurions Correlation**

إذا كانت العلاقة بين ظاهرتين نتيجة وجود عامل مشترك يؤثر في الظاهرتين.

"إذا درسنا العلاقة بين طول قدم الطفل ودرجة ذكائه، لوجدنا إن هناك ميل كبير للظاهرتين للتغير معاً، أي أن هناك ارتباط قوى، ولكن هذا غير صحيح، حيث إننا أهملنا العامل المشترك الذي يؤثر في الظاهرتين معاً وهو عمر الطفل، فكلما زاد عمره ازداد طول قومه وذكائه، لذلك يسمى هذا النوع بالارتباط الوهمي".

4/7 الارتبــاط البسيط Simple Correlation

إذا كانت العلاقة بين متغيرين فقط يكون الارتباط ارتباط بسيط مثل قياس العلاقة بين حجم الطلب على سلعة معينة وسعرها، أو العلاقة بين حجم الدخل القومي وعدد السكان.

5/7 الارتباط المتعدد Multiple Correlation

عند دراسة العلاقة بين أكثر من ظاهرتين، أي بين التغير في عدد من المتغيرات، والتغير في المتغير محل الدراسة، فيطلق على الارتباط في هذه الحالة الارتباط المتعدد Multiple Correlation مثل سعر السلعة ودخل المستهلك واسعار السلع الأخرى وأثر كل هذه الظواهر من حجم الطلب على سلعة معينة.

6/7 الارتباط الجزئي Partial Correlation

عندما يلجأ الباحث إلى إبقاء عدة ظواهر ثابتة، ودراسة أثر ظاهرة واحدة على المتغير محل الدراسة فيكون الارتباط في هذه الحالة ارتباط جزئي Partial correlation بالنسبة إلى كيفية قياس الارتباط البسيط، نلاحظ أن:

- العلاقة بين المتغيرين محل الدراسة، أما أن تكون خطية Linear، أي أن التغير في أحد المتغيرات يؤدي إلى تغير الآخر، بنسبة ثابتة.
- أو قد تكون العلاقة غير خطية Non-Linear أي أن التغير في أحدهما يؤدي إلى تغير الآخر بنسب متغيرة.

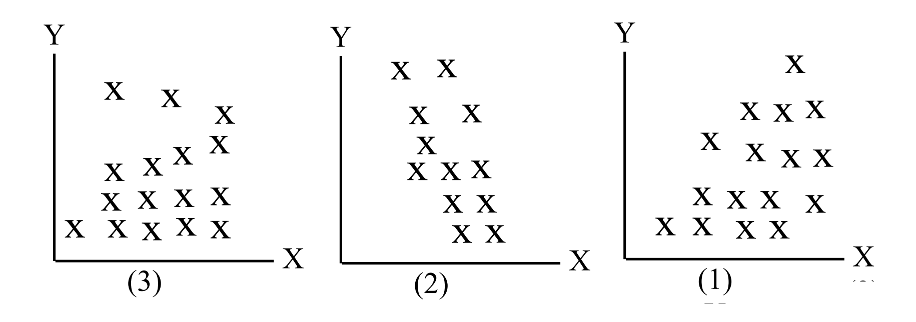
ويلاحظ: أن درجة الارتباط التي نحسبها، تعتمد على نوع العلاقـة المفترضـة ومـدى صحة هـذا الافتراض، وهذه حقيقة يجب مراعاتها عند تفسير النتائج التي نحصل عليها.

- فإذا كانت العلاقة بين المتغيرين خطية، وحسبنا درجة الارتبـاط فكانـت منخفضـة، فـإن هـذا لا يعني أكثر من عدم وجود علاقة خطية بين المتغيرين، ولا يعني عدم وجود ارتباط بينهما.
- وإذا كانت العلاقة بين المتغـيرين غـير خطيـة، فقـد ترتفـع درجـة الارتبـاطن مـما يـدل عـلى أن المنحنى أفضل من الخط المستقيم في وصف العلاقة بين المتغيرين.
- ومن وجهة أخرى، فإن وجود ارتباط قوي طبقاً للمقاييس المستخدمة قد لا يعني وجـود علاقـة أصلاً بين المتغيرين، كما في حالة الارتباط الوهمي Dummy Correlation.

7/7 قياس الارتباط بين بيانات غير مبوبـــة

1/7/7 شكل الانتشار **Scatter Diagram**

بعد جمع بيانات عن المتغيرين (Y,X) محل الدراسة، نقوم برصد أزواج القـيم عـلى رسـم بيـاني، فنحصل على شكل الانتشار الذي قد يأخذ واحدة من عدة صور، كما هو موضح.



شكل الانتشار

يفيد شكل الانتشار في إعطاء صورة عن قوة العلاقة بين المتغيرين ونوعها.

- ففي الشكل (1) نلاحظ أن المتغيرين يتغيران في اتجاه واحد، فقيمة مرتفعة للمتغير X يتبعها قيمة مرتفعة للمتغير Y، مما يدل على وجود علاقة طردية بين المتغيرين، أو أن الارتباط موجب Positive Correlation أي أن تغير أحد المتغيرين يؤدي إلى تغير الآخر في نفس الاتجاه.
- أما في (2) فنلاحظ أن هناك ارتباط سلبي Negative أي أن التغير في ظاهرة يصاحبه تغير في الاتجاه المضاد في الظاهرة الأخرى، حيث أن القيم الكبيرة على المحور الأفقي تصاحب القيم الصغيرة على المحور الرأسي، والعكس بالعكس.
- أما في الشكل (3): فأحياناً يصاحب القيم الكبيرة للمتغير X قيم كبيرة للمتغير Y، وأحياناً أخرى تصاحبها قيم صغيرة، مما يدل على عدم وجود ارتباط بين المتغيرين.

2/7/7 حساب معامل الارتباط Corrlation Coefficient

الارتباط يدرس ميل ظاهرتين للتغير معاً.

يحسب كارل بيرسون Corl Person المتوسط الحسابي لحواصل ضرب كل زوج من المشاهدات معبراً عنها بوحدات معيارية كمقياس لدرجة الارتباط الخطي بين الظاهرتين. حيث أنه لو كانت العلاقة بين الظاهرتين غير خطية، فهذه الطريقة لا تصلح كمقياس للارتباط.

لذا يفضل استخدام شكل الانتشار وكخطوة أولى للتأكد من طبيعة العلاقة قبل تطبيق معامل بيرسون لقياس مدى قوتها.

وهناك أيضاً مقياس آخر يستعمل رتب القيم بدلاً من قيمها الفعلية لذا يسمى بمعامل الرتب أو معامل سبيرمان.

معامل ارتباط بيرسون Person Foefficent (r)

$$r = \frac{1}{N}\left[\frac{(X_1 - \overline{X})}{Sx} . \frac{Y - \overline{Y}}{Sy} + \frac{X_2 - \overline{X}}{Sy} . + ... + \frac{Xn - \overline{X}}{Sx} \frac{Yn - \overline{Y}}{Sy} \right]$$

حيث Sx الانحراف المعياري للمتغير Sy, x الانحراف المعياري للمتغيرY، $\overline{Y}$ ، الوسط الحسابي للمتغير $\overline{X}$, Y الوسط الحسابي للمتغير X وعدد المشاهدات هو N.

- إذا كان هناك ارتباط كبير بين الظاهرتين، فالتغير في أحدهما سيؤدي إلى تغير في الأخرى، فيكون حاصل الضرب كبيراً، وبالتالي المتوسط الحسابي.
- وعندما يكون الارتباط موجب، فبتغير X سيؤدي إلى تغير Y في نفس الاتجاه، وبالتالي يكون حاصل الضرب موجب للتشابه.
- أما في حالة الارتباط العكس، فإن الإشارتين سيكونا مختلفتين وبالتالي حاصل الضرب سالب.
- لو كان الارتباط ضعيف بحيث أن بعض الانحرافات موجبة والأخرى سالبة، فتنخفض قيمة المعامل، وتتراوح قيمة معامل الارتباط بين -1 ، -1.
- فإذا كان مساوياً +1، فهذا يدل على أن التغير في X يشرح كل التغير في Y، مما يدل على وجود ارتباط تام بني التغير في الظاهرتين.
- أما إذا كانت العلاقة سالبة، والتغير في Y. يعود بالكامل إلى التغير في X، فإن معامل الارتباط يساوي -1، مما يدل على وجود ارتباط تام أيضاً.
- في الحالتين، جميع النقاط في شكل الانتشار ستقع على خط مستقيم، إما صاعد في الحالة الأولى، أو نازل في الثانية.
- من النادر أن نحصل على ارتباط تام بين متغيرين، خاصة عند دراسة الظواهر الاجتماعية نتيجة وجود متغيرات أخرى لم نأخذها في الحسبان، لذا تقل قيمة المعامل عن الواحد الصحيح، مع ملاحظة أنه كلما قربت القمية المطلقة للمعامل من الواحد الصحيح كلما دل ذلك على قوة الارتباط بين المتغيرين، علماً بأن الإشارة توضع طبعة العلاقة أما موجبة أو سالبة ولا تعكس درجة الارتباط أو قوته.

وعلى ذلك فإن معامل ارتباط بيرسون يعبر عنه بالمعادلة التالية في صورته الأصلية.

$$r = \frac{\sum Xy}{NSXSy}$$

حيث X هي انحرافات x عن وسطها الحسابي.

Y هي انحرافات y عن وسطها الحسابي.

مثال:

حصل باحث على قراءات عن حجم الأسرة (y) وحجم إنفاقها (x)، والمطلوب حساب درجة ونوع الارتباط بين المتغيرين:

X:	1	3	4	6	8	9	11	14
Y:	1	2	4	4	5	7	8	9

الحل:

1	2	3	4	5	6	7
X	**Y**	**X**	**Y**	**XY**	X^2	Y^2
1	**1**	**-6**	**-4**	**24**	**36**	**16**
3	**2**	**-4**	**-3**	**12**	**16**	**9**
4	**4**	**-3**	**-1**	**3**	**9**	**1**
6	**4**	**-1**	**-1**	**1**	**1**	**1**
8	**5**	**1**	**0**	**0**	**1**	**0**
9	**7**	**2**	**2**	**4**	**4**	**4**
11	**8**	**4**	**3**	**12**	**16**	**9**
14	**9**	**7**	**4**	**28**	**44**	**16**
$\sum$ **56**	**40**	**0**	**0**	**84**	**132**	**56**
				$\sum xdyd$	$\sum xd^2$	$\sum yd^2$

حيث أننا حسبنا المتوسط الحسابي X عن طريق $X = \frac{\sum X}{N}$.

ومنها حسبنا ا نحرافات X عن وسطها الحسابي في الخانة 3.

وبتكرار الخطوات بالنسبة للمتغير Y نستطيع حساب الخانة 4 التي تعبر عن $\left(Y - \overline{Y}\right)$. الخانة الخامسة هي حاصل ضرب الخانتين السابقتين.

أمّا الخانة رقم (6) فهي عبارة عن مربعات انحرافات X عن وسطها الحسابي، أي $\left(X - \overline{X}\right)^2$ أي مربع القيم الواردة في الخانة (3). والخانة الأخيرة هي مربعات الانحرافات (Y) عن وسطها الحسابي.

نحتاج إلى حساب الانحراف المعياري للمتغيرين، ويتم عن طريق التعويض في القانون:

$$Sx = \sqrt{\frac{\sum x^2}{N}} \quad \sqrt{\frac{132}{8}} = 4.06$$

$$Sy = \sqrt{\frac{\sum y^2}{N}} \quad \sqrt{\frac{56}{8}} = 2.65$$

وبالتطبيق في الصورة العامة لحساب معامل الارتباطن نجد أن:

$$r = \frac{84}{(8)\,(2.65)\,(4.06)} = 0.975$$

أي أن الارتباط بين المتغير (x) والمتغير (y) قوي، حيث أن قيمة المعامل قريبة من الواحد الصحيح، وأن العلاقة بينهما طردية، لأن الإشارة موجبة.

معامل ارتباط بيرسون

$$r = \frac{\sum X'Y' - \frac{\sum X' \sum Y'}{N}}{NSx\ SY}$$

حيث تعبر X' عن انحرافات X عن وسط فرضي ، و Y' عن انحرافات Y عـن وسـط فـرضي، وقد يكون الوسط الفرضي للمتغيرين قيمة واحدة، أو قيمتـين مختلفتـين، ويفضـل أن يكونـا مـن قـيم المشاهدات لتبسيط العمليات الحسابية أكثر.

كما يمكن حساب الانحـراف المعيـاري باسـتخدام الانحرافـات عـن الوسـط الفـرضي المسـتخدمة باستخدام القانون التالي:

$$SX = \sqrt{\frac{\sum X'^2}{N} - \left(\frac{\sum X'}{N}\right)^2}$$

$$Sy = \sqrt{\frac{\sum X'^2}{N} - \left(\frac{\sum Y'}{N}\right)^2}$$

ولتوضيح حساب معامـل الارتبـاط، يؤخـذ انحرافـات عـن وسـط فـرضي "بحـل المثـال السـابق" وحسـاب انحرافـات (X) عـن القيمـة 8، وانحرافـات (Y) عـن القيمـة 7، وحسـاب الخانـات الـواردة في الجدول التالي:

1	2	3	4	5	6	7
X	Y	X'	Y'	$X'Y'$	$(X')^2$	$(Y')^2$
1	1	-7	-6	42	49	16
3	2	-5	-5	25	25	25
4	4	-4	-3	12	16	9
6	4	-2	-3	6	4	9
8	5	0	-2	0	0	4
9	7	1	0	0	1	0
11	8	3	1	3	9	1
14	9	6	2	12	36	4
$\sum$ 56	40	-8	-16	100	140	88

$$Sy = \sqrt{\frac{140}{8} - \left(\frac{-8}{8}\right)^2} = 4.06$$

$$Sy = \sqrt{\frac{88}{8} - \left(\frac{-16}{8}\right)^2} = 2.65$$

وبالتعويض في القانون نحصل على قيمة معامل الارتباط، كالآتي:

$$r = \frac{100 - \frac{(-8)(-16)}{(8)}}{(8)(4.06)(2.65)} = 0.975$$

وهي نفس النتيجة السابقة بالطبع، حيث أننا لم نغير قيمة الظاهرتين، بل استخدمنا صورة أخرى من معامل بيرسون لتبسيط العمليات الحسابية.

ويمكن الحصول على معامل بيرسون باستخدام الأرقام الخام مباشرة، خاصة عندما تكون المشاهدات قليلة، وقيمتها منخفضة، وتطبق الصيغة التالية من القانون:

$$r = \frac{\sum XY - NX'Y'}{NSxSy}$$

ويمكن حساب الانحراف المعياري باستخدام الأرقام الخام مباشرة بتطبيق القانون:

$$Sx = \sqrt{\frac{\sum X^2}{N} - \left(\frac{\sum X}{N}\right)^2}$$

$$Sy = \sqrt{\frac{\sum Y^2}{N} - \left(\frac{\sum Y}{N}\right)^2}$$

وبتطبيق هذه الصيغة على نفس المثال المستخدم، نحتاج على حساب البيانات الواردة في الجدول التالي:

X	Y	XY	X^2	Y^2
1	1	1	1	1
3	2	6	6	4
4	4	16	16	16
6	4	24	24	16
8	5	40	40	25
9	7	63	63	49
11	8	88	88	64
14	9	126	126	81
$\sum$ 56	40	364	524	256

$$Sx = \sqrt{\frac{524}{8} - \left(\frac{56}{8}\right)^2} = 43.06$$

$$Sy = \sqrt{\frac{256}{8} - \left(\frac{40}{8}\right)^2} = 2.65$$

ونحتاج إلى حساب المتوسطات الحسابية

$$\overline{X} = \frac{56}{8} = 7$$

$$\overline{Y} = \frac{40}{8} = 5$$

$$r = \frac{364 - (8)\,(7)\,(5)}{(8)\,(4.06)\,(2.65)} = 0.975$$

وهي نفس النتيجة السابقة أيضاً، وتدل على ارتباط موجب قوي "ارتباط موجب قوي" بين الظاهرتين.

8/7 معامل سبيرمان (ارتباط الرتب) **Spearman**

اشتق (سبيرمان) القانون التالي كحالة خاصة من قانون (بيرسون) حيث استخدم (رتب القيم) بدلاً من القيم نفسها.

ويحسب معامل الارتباط "معامل سبيرمان" بالقانون التالي:

$$r = 1 - \frac{6\sum D^2}{N(N^2 - 1)}$$

حيث D هي الفروق Deviation الفروق المطلقة بين رتب المتغيرين.

N عدد المفردات.

ولتطبيق هذا القانون، نقوم بإعطاء رتب لقيم المتغيرين كما هي واردة ضمن المشاهدات دون تغيير، مع ضرورة استخدام نفس الترتيب (تصاعدي أو تنازلي) بالنسبة للمتغيرين.

لا يجوز استخدام ترتيب تصاعدي للمتغير الأول، وتنازلي للثاني، وبالعكس.

نأخذ الفروق المطلقة بين الرتب (والذي يمثل التقارب أو عدم التقارب) بين قيم المتغيرين، ثم نربعها (أي الفروق بين الرتب)، ونعوض بالقانون.

مثال:

احسب معامل ارتباط الرتب (معامل سبيرمان) بين المتغيرين، في الجدول:

X:	5	7	9	10	8	6
Y:	1	3	2	6	5	4

الحل:

سنستعمل ترتيب تنازلي، حيث نأخذ أكبر قيمة الرتبة 1 ثم نأخذ القيمة الثانية الرتبة 2 وهكذا ونحسب الفروق في الرتب، كما في الجدول:

X	Y	رتب X	رتب Y	D	D^2
5	1	6	6	0	0
7	3	4	4	0	0
9	2	2	5	3	9
10	6	1	1	0	0
8	5	3	2	1	1
6	4	5	3	2	4
					14

$$r = 1 - \frac{(6)\,(14)}{(6)\,(36-1)} = 1 - 0.4 = 0.6$$

يمتاز معامل سبيرمان (ارتباط الرتب)، بسهولة حساب، غلا أنه يعطي قيمة تقريبية أقل دقة من معامل بيرسون، وذلك لأنه يعتمد على تراكيب تراتيب القيم، بنيما معامل بيرسون يعتمد على القيم نفسها.

ومن ناحية أخرى، يتيح معامل ارتباط سبيرمان، إمكانية دراسة الارتباط بين بيانات وصفية سواءً للظاهرتين محل الدراسة، أو لأحدهما، الأمر الذي لم يكن ممكناً

عند استخدام معامل بيرسون، ويتم ذلك بإعطاء رتب لقيم الظاهرتين، وتطبيق القانون السابق.

مثال:

احسب درجة الارتباط بين درجات الطلبة في مادة الإحصاء، ودرجة ذكائهم، علماً بأن العينة التي تم الحصول عليها، هي:

تحت المتوسط	فوق المتوسط	عالي	ضعيف جداً	ضعيف	متوسط	عالي	عالي	الذكاء(x)
76	89	90	52	59	90	95	90	درجة الإحصاء (y)

الحل:

يلاحظ أنه عند تكرار نفس القيمة أو التقدير فإنهم يأخذوا نفس الترتيب. ويتم حساب ذلك الترتيب، بأخذ المتوسط الحسابي للرتب المفروض أن تعطي للقيمة المتكررة.

فمثلاً: درجة الذكاء "عالي" سوف تأخذ الرتب 1m2m3 في حالة استخدام ترتيب تنازلي، حيث أنها مكرر ة ثلاثة مرات، لكنهم أعطو الترتيب المتوسط، أي مجموع الرتب مقسم ÷ على عددهم.

$$\frac{1+2+3}{3}=2$$

وهكذا أيضاً بالنسبة للدرجة 90، حيث أنها تكررت أكثر من مرة وبالتالي يحسب لها ترتيب هو عبارة عن متوسط الرتب المفروض أن تأخذها، وهي في مثالنا هذا:

$$\frac{2+3+4}{3}=3$$

وبالتالي يمكن أن تكون الرتب كسوراً عشرية، ويتم حساب الفروق كما في الجدول التالي:

$$r = \frac{6\sum D^2}{N(N^2 - 1)}$$

X	Y	رتب X	رتب Y	D	D^2
عالي	90	2	3	1	1
عالي	95	2	1	1	1
متوسط	90	5	3	2	4
ضعيف	59	7	7	0	0
ضعيف جداً	52	8	8	0	0
عالي	90	2	3	1	1
فوق المتوسط	89	4	5	1	1
تحت المتوسط	75	6	6	0	0
					8

وبالتعويض في القانون:

$$r = 1 - \frac{(6)(8)}{8(64 - 1)}$$

$$= 1 - 0.095 = 0.905$$

9/7 قياس الارتباط للبيانات المبوبـــة

إن زيادة عدد المشاهدات يستوجب تبويبها وعرضها في جداول تكرارية، لـذا لا بـد مـن التفكـير في كيفية حساب الارتباط للجداول التكرارية.

توجد طريقتين:

1/9/7 معامل ارتباط بيرسون

يعتمد هذا المعامل على حساب الانحرافات عن الوسط الحسابي، ثـم التعبـير عنهـا كنسـبة مـن الانحراف المعياري، لاستبعاد أثـر وحـدات القيـاس ثـم أخـذ المتوسـط الحسـابي لحاصـل ضرب الأزواج المتناظرة.

توجد عدة صيغ منها:

1- استخدام الانحرافات عن الوسط الحسابي

احسب معامل ارتباط بيرسون بأخذ الانحرافات عن الوسط الحسابي من الجدول التكراري التالي:

فئات X / فئات Y	20-	30-	-40	50-	60-70	5
10-	1	2	2	-	-	5
14 -	3	8	10	4	-	25
18 -	1	7	12	11	4	35
22 - 26	-	3	6	5	1	15
fx	5	20	30	20	5	80

ويتم حساب معامل الارتباط، باستخدام القانون:

$$r = \frac{\sum Fxy}{NSxSY}$$

حيث:

N = مجموع التكرارات

F = تكرارات الخانات الداخلية للجدول.

X = انحرافات X عن وسطها الحسابي.

Y = انحرافات Y عن وسطها الحسابي.

Sx = الانحراف المعياري للمتغير X.

Sy = الانحراف المعياري للمتغير Y.

الحل: لإيجاد الحل نتبع الخطوات التالية:

نحسب المتوسط الحسابي لكل من المتغيرين، وكما نعلم فإن:

$$\overline{Y} = \frac{\sum FyY}{N}$$

حيث Y هي مركز الفئة

ويكون المتوسط الحسابي للمتغير الأول Y هو:

$$\overline{Y} = \frac{12\times5+16\times25+20\times35+24\times15}{80} = 19$$

وبالمثل المتوسط الحسابي للمتغير X، ونجد أنه:

$$\overline{X} = \frac{25\times5+35\times20+45\times30+55\times20+65\times5}{80} = 45$$

- نحسب انحرافات مراكز الفسئات عن الوسط الحسابي لكل متغير (x,y).
- نضرب الانحرافات X في التكرارات (fxx^2, fyy^2).
- من البيانات السابقة، نحسب الانحراف المعياري لكل من المتغيرين ونلخص هذه الخطوات، في الجدول التالي.

				Fxx^2	**2000**	**2000**	**0**	**2000**	**2000**	**2000**
				Fxx	-100	-200	0	200	100	0
				X	-20	-10	0	10		
				X	25	35	45	55		
Fyy^2	**fyy**	**y**	**Y**		**20-**	**30-**	**40-**	**50-**	**60-**	
245	-35	-7	12	10-	1	2	2	-	-	2
225	-75	-3	16	14-	3	8	10	4	-	25
35	35	1	20	18-	1	7	12	11	4	35
375	75	5	24	22-	-	3	6	5	1	15
880	**0**				**5**	**20**	**30**	**20**	**5**	**80**

ويتم حساب الانحراف المعياري، لكل متغير باستخدام القوانين.

$$Sx = \sqrt{\frac{\sum fxx^2}{fx}} = \sqrt{\frac{8000}{80}} = 10$$

$$Sx = \sqrt{\frac{\sum fyy^2}{fy}} = \sqrt{\frac{800}{80}} = 3.3$$

للحصول على البسط في الصورة المستخدمة من قانون معامل الارتباط السابق، نحتاج إلى خطوة أخرى، حيث:

نضرب التكرارات الداخلية للجدول في كل من انحرافات X عـن وسـطها الحسـابي، وانحرافـات y عن وسطها الحسابي، كما يتضح من الجدول التالي:

X / Y	20-	-10	0	10	20	$\sum$
-7	1 (140)	2(140)	2 (0)	-(0)	- (0)	(280)
-3	3 (180)	8(240)	10 (0)	4 (-120)	- (0)	(300)
1	1 (-20)	7(-70)	12 (0)	11 (110)	4(80)	(100)
5		3 (-150)	6 (0)	5 (250)	(100)	(200)
$\sum$	**(300)**	**(160)**	**0**	**(240)**	**(180)**	**880**

بجانب كل تكرار وضعنا حاصل ضرب الأرقـام الثلاثـة y, x والتكـرار الـداخلي في الخانـة، فـالرقم (140) هو حاصل ضرب (7-) (1) (20-)، أما الرقم الثاني في العمود الأول (180) فهو حاصـل ضرب -) (20-) (3) (3، وهكذا بجانب كل تكرار نكتب حاصل ضرب الأرقام الثلاثة، مع مراعاة الإشارة.

نجمع كافة الأرقام الناتجة عـن حاصـل الضرب (أي الأرقـام بـين القوسـين)، نحصـل عـلى قيمـة البسط $\sum$FXY، ويجب أن يتساوى حاصـل الجمـع أفقيـاً ورأسـياً لتلـك الأرقـام، وإلا فيجـب مراجعـة حواصل الضرب.

نطبق القانون (20-) أما الرقم الثاني

$$r = \frac{880}{(80)(10)(33)} = 0.33$$

"أي أن الارتباط بين المتغيرين طردي وغير قوي نسبياً".

10/7 استخدام الانحرافات عن وسط فرضي

يمكن حساب الانحرافات عن وسط فرضي، وليكن أحد مراكز الفئات بالنسبة للمتغيرين. لتبسيط العملايت الحسابية:

- حيث لا نحسب المتوسط الحسابي، والذي قد يحتوي على كسر عشري.
- قد يختلف الوسط الفرضي للمتغيرين، أو نأخذ الانحرافات عن نفس الوسط الفرضي.
- كذلك، يمكن القسمة على عامل مشترك لتبسيط الحساب أكثر واكثر.

سنستخدم الصورة الثالية للقانون:

$$r = \frac{\sum FX^*Y^* \frac{\left(\sum FxX^*\right)\left(\sum fyY^*\right)}{N}}{NSx\ Sy}$$

حيث تمثل X^* انحرافات X عن وسط فرضي.

وتمثل Y^* انحرافات Y عن وسط فرضي.

أما بقية المتغيرات، فلها نفس المعنى السابق.

مثـــال:

احسب معامل بيرسون من الجدول التالي، بأخذ انحرافات عن وسط فرضي:

Y \ X	10-	14-	18-	22-26	fy
10-	6	4	-	2	12
20-	-	8	20	-	28
30-	6	12	12	6	36
40-50	-	4	4	16	24
fx	**12**	**28**	**36**	**24**	**100**

الحل:

- نأخذ انحرافات X عن أحد مراكز الفئات، وليكن 20.
- نقسم الانحرافات على العامل المشترك 4 تبسيط العمليات الحسابية أكثر.

- نأخذ انحرافات Y عن مركز أحد الفئات، وليكن 35، ونقسم على العامل المشترك 10.
- نضرب الانحرافات X في التكرار المقابل.
- نربع الانحرافات ونضربها في التكرار.
- حساب الانحراف المعياري.

هذه الخطوات موضحة في الجدول التالي:

				Fxx^2	48	28	0	24	90
				Fxx'	-24	-28	0	24	-28
				X'	-2	-1	0	1	
				X	12	16	20	24	
Fyy'^2 48	**fy' -24**	**Y'**	**Y**						**fy**
48	-24	-2	15	10-	6	4	-	2	12
28	-28	-1	25	20-	-	8	20	-	28
0	0	0	35	30-	6	12	12	6	36
24	24	1	45	40-	-	4	4	16	24
90	**-28**			**fx**	**12**	**28**	**36**	**24**	**100**

$$Sx = L\sqrt{\frac{\sum fxx'^2}{fx} - \left(\frac{\sum fxx}{fx}\right)^2} = 4\sqrt{\frac{90}{100} - \left(\frac{-28}{100}\right)^2}$$

$$= \boxed{3.625}$$

$$Sy = L'\sqrt{\frac{\sum fyy^2}{fy} - \left(\frac{\sum fyy}{fy}\right)^2} = 10\sqrt{\frac{90}{100} - \left(\frac{-29}{100}\right)^2}$$

$$= \boxed{9.064}$$

لحساب البسط نحتاج إلى جدول آخر لضرب انحرافات X في انحرافات Y في التكرارات الداخلية، وهذا ما سيتم توضيحه في الجدول التالي:

Y′ \ X′	-2	-1	0	1	
-2	6(24)	4(8)	-(0)	2(-4)	(28)
-1	-(0)	8(8)	20(0)	-(0)	(8)
0	6(0)	12(0)	12(0)	6(0)	(0)
1	(-0)	4(-1)	4(0)	16(16)	(12)
	(24)	**(12)**	**(0)**	**(12)**	**48**

نطبق القانون بعد إدخال المعاملات المشتركة في الحسبان:

$$r = (4)\,(10)\left[\frac{48 - \frac{(-28)\,(-28)}{100}}{(100)(9.064)(3.625)}\right] = .489$$

1/10/7 استخدام الأرقام الخام مباشرة

عند استخدام هذا الأسلوب لحساب معامل بيرسون، نستخدم الصور التالية لحساب الانحراف المعياري، ومعامل الارتباط:

$$r = \frac{\sum FXY - NX'Y'}{NSxSy}$$

$$Sx = \sqrt{\frac{\sum FxX^2}{Fx} - \left(\frac{\sum FxX}{Fx}\right)^2}$$

$$Sy = \sqrt{\frac{\sum FyY^2}{Fy} - \left(\frac{\sum FyY}{Fy}\right)^2}$$

مثال:

احسب باستخدام الأرقام الخام، معامل الارتباط من الجدول التكراري المعطى في المثال الأسبق.

الحل: جد أن الحل يشممل الخطوات التالية:

نحسب المتوسط الحسابي لكل من المتغيرين بنفس الأسلوب، ومنها نعلم أن

$$X' = 45 \quad , \quad Y' = 19$$

حساب الانحراف المعياري للمتغيرين، كالآتي:

X	FxX	FxX^2	Y	FyY	FyY^2
25	125	3125	12	60	720
35	700	24500	16	400	6400
45	1350	60750	20	700	14000
55	1100	60600	24	360	8640
65	325	21125			
	3600	170.000		1520	29760

$$Sy = \sqrt{\frac{29760}{80} - \left(\frac{1520}{80}\right)^2} = 3.3$$

$$Sx = \sqrt{\frac{170.000}{80} - \left(\frac{3600}{80}\right)^2} = 10$$

حساب معامل الارتباط، ويتم بضرب مراكز الفئات (أي المتغـيرين) في التكـرارات الداخليـة، كـما موضح في الجدول.

Y \ X	25	35	45	55	65	$\sum$
12	1(300)	2(840)	2(1080)	-(0)	(-0)	(2220)
16	3(1200)	8(4480)	10(7200)	4 (3520)	(-0)	(16400)
20	1(500)	7(4900)	12(10800)	11(12100)	4(5200)	(23500)
24	-(0)	3(2520)	86(6480)	5(660)	1(1560)	(17160)
$\sum$	**(2000)**	**(12740)**	**(25560)**	**(22220)**	**(6760)**	**69280**

فعلى سبيل المثال، الرقم الأول (300)، هو عبارة عن حاصل ضرب (1) (25) (12)، وهكـذا نحسـب بقية الأرقام الواردة بين قوسين، ثم نجمع أفقياً ورأسياً للمراجعة.

بالتطبيق في القانون:

$$r = \frac{69280 - [80(45)(14)]}{(80)(10)(3.3)} = 0.33$$

2/10/7 معامل ارتباط الرتب:

هناك عدة طرق لحساب معامل ارتباط الرتب (سبيرمان) للجداول التكرارية، منها طريقة أقطار المجاميع المتساوية، وطريقة أقطار الفروق المتساوية.ت وسنبدأ بالطريقة الأخيرة، والتي تتطلب قبل تطبيقها التأكد من أن فئات كل من المتغيرين متساوية في العدد، وأنها أي الفئات، مرتبة تصاعدياً في الجدول التكراري.

سنوضح الخطوات المطلوبة في طريقة أقطار الفروق المتساوية باستخدام المثال التالي:

مثال:

احسب معامل ارتباط الرتب من الجدول التكراري التالي:

Y \ X	10-	20-	30	40-	50-60-	fy
35-	10	4	-	-	-	14
50-	20	22	8	2	-	52
65-	10	28	34	18	6	96
80-	2	18	32	22	10	84
95-110	2	6	24	14	8	54
fx	**44**	**78**	**98**	**56**	**24**	**300**

في هذه الطريقة نحتاج إلى حساب كلام التباين والانحراف المعياري للظاهرتين، فنقوم بإعطائهما ترتيب إما تصاعدي أو تنازلي، شريطة تطبيق نفس النظام على الظاهرتين، ثم نأخذ الانحرافات عن واحدة من هذه الرتب، وقد تكون نفس الرتبة أو مختلفة بالنسبة للظاهرتين. ثم نضرب الانحرافات في التكرارات، ثم مربع الانحرافات في X التكرارات لنصل إلى المطلوب كما في الشكل التالي:

رتب Y	Y	Fy	fyF	fyY^2	رتب X	X	Fx	fxX	fxX^2
1	-2	14	-28	56	1	-2	44	-88	176
2	-1	52	-52	52	2	-1	78	-78	78
3	0	96	0	0	3	0	98	0	0
4	1	84	84	84	4	1	56	56	56
5	2	54	108	216	5	2	24	48	96
			112	**408**				**-68**	**406**

$$\sigma y = \frac{\sum FyY^2}{fy} - \left(\frac{\sum FyY}{Fy}\right)^2 \qquad \sigma y = \frac{\sum FxX^2}{fx} - \left(\frac{\sum FxX}{Fx}\right)^2$$

$$\sigma y = \frac{408}{300} - \left(\frac{112}{300}\right)^2 \qquad \sigma y = \frac{406}{300} - \left(\frac{-62}{300}\right)^2$$

= 1.21   1.30

وحيث أن الانحراف المعياري هو الجذر التربيعي للتباين، فإن:

$$Sy = \sqrt{1.21} = 1.10 \qquad Sy = \sqrt{1.3} = 1.14$$

تحتاج هذه الطريقة أيضاً إلى حساب تباين الفروق σD الذي يتم باستخدام الجدول التالي.

3/10/7 الارتباط بين الظواهر الوصفية:

كثيراً ما تتطلب الدراسة والبحث في مجال الدراسات الاجتماعية، دراسة درجة العلاقة بين ظواهر غير كمية، لا يمكن التعبير عنها بالأرقام، بل توصف فقط مثل، الارتباط بين الجنس (ذكر أم أنثى)، وبين الحالة التعليمية لمجموعة من الأفراد، أو دراسة الارتباط بين التطعيم بمصل واقي والإصابة بمرض معين، وغير ذلك من الأمثلة.

في هذه الحالات، يمكننا استخدام معامل الاقتران أو معامل التوافق، طبقاً لطبيعة الظواهر محل الدراسة وعدد الأنواع التي تنقسم إليها كل ظاهرة.

معامل الاقتران:

وقد وضعه ج، أ.بول لدراسة العلاقة بين ظاهرتين وصيفيتين، تنقسم كل منهما على نوعين أو مجموعتين فقط، كالآتي:

مجموعة 2 للظاهرة الأولى	مجموعة 1 للظاهرة الأولى	
b	d	مجموعة 1 للظاهرة الثانية
d	c	مجموعة 2 للظاهرة الثانية

ويحسب معامل الاقتران بين هذين الصفتين، بالقانون:

$$tA = \frac{ad - bc}{ad + bc}$$

ويأخذ هذا المعامل القيم بين $-1 \leq tA \leq 1$

وكلما اقتربت قيمته من الصفر، كان ذلك دليلاً على عدم وجود اقتران بين الظاهرتين محل الدراسة.

أما إذا كانت قيمة المعامل مساوية (الواحد الصحيح)

(وهذا يتطلب إن bc يساوي الصفر)، فإن هناك علاقة تامة بين الظاهرتين محل الدراسة.

مثال:

في دراسة طبية لمصل جديد ضد مرض ما على عينة من الأفراد توفرت لدينا البيانات التالية:

	يطعموا	لم يطعموا	المجموع
أصيب بالمرض	78	131	309
لم يصب بالمرض	318	52	370
المجموع	**396**	**183**	**579**

$$\therefore tA = \frac{78 \times 52 - 318 \times 131}{78 \times 52 + 318 \times 131} = \frac{4056 - 41658}{3} = -.82$$

"وتوضع النتيجة أن هناك علاقة قوية بني التطعيم وعدم الإصابة بالمرض".

4/10/7 معامل التوافق

وضعه بيرسون لدراسة العلاقة بين الظواهر الوصفية التي تنقسم إلى أكثر من نوعين، ففي هذه الحالة لا يساعدنا معامل الاقتران السالب الذكر، ويستخدم معامل التوافق الذي يصلح أيضاً لقياس العلاقة بين ظواهر كمية قابلة للقياس، وأخرى وصفية لا يمكن قياسها.

يحسب معامل التوافق من القانون التالي:

$$tp = \sqrt{\frac{G - 1}{G}}$$

حيث (G) هي مجموع خارج قسمة مربع كـل تكـرار في X الجـدول التكـراري ÷ عـلى مجمـوع بيانات الصف والعمود التي يقع فيها هذا التكرار.

أي أنه لتطبيق هذا القانون نقوم بتربيع كل تكـرار وارد في الجـدول التكـراري ونقسـمه ÷ عـلى حاصر ضرب التكرار الكلي الرأسي في x التكرار الأفقي، كما في المثال التالي:

مثـــال:

احسب درجة الارتباط بين تقدير الطالب في المادة (X) والمادة (Y) من البيانات التالية:

X / Y	مقبول	جيد	جيد جداً	ممتاز	المجموع
مقبول	10	18	23	14	65
جيد	45	103	82	50	280
جيد جداً	27	50	30	21	128
المجموع	**82**	**171**	**135**	**85**	**473**

$$G = \frac{(10)2}{82 \times 65} + \frac{(45)^2}{82 \times 280} + \frac{(27)^2}{82 \times 128} + \frac{(18)^2}{171 \times 280} + \frac{(103)^2}{171 \times 250} + \frac{(23)^2}{135 \times 65}$$

$$+ \frac{(82)^2}{135 \times 280} + \frac{(30)^2}{135 \times 128} + \frac{(14)^2}{85 \times 65} + \frac{(50)^2}{85 \times 280} + \frac{(21)^2}{85128} = 1.0066$$

$$\therefore tp = \sqrt{\frac{1.0066 - 1}{1.0066}} = 0.08$$

"أي أن هناك ارتباط ضعيف جداً بين تقديرات الطالب في المادتين".

الوحدة الثامنة

الارتباط والانحدار على الحاسب الإلكتروني

1/8 تطبيقات الارتباط على الحاسوب **Correlation**

اختصت جميع الحالات السابقة بمعالجة متغير واحد، فالوسط الحسابي، أو الانحـراف المعيـاري، أو إحصائية الاختبار ويخص كلا منها متغير واحد لأن التوزيع لا ينتج لأكثر من متغير واحد.

أما إذا كانت لكل قيمة من المتغير (سر)قيمة أخرى تناظرها بمتغـير آخـر (صر)، فالبيانـات ذات الأزواج المرتبة تسمى بالبيانات ذات البعدين، ويسمى كل منها بالمتغير العشوائي ذي البعـدين، ويتبـع كل متغير من المتغيرين توزيعاً خاصاص يسمى بالتوزيع الهامش للمتغير.

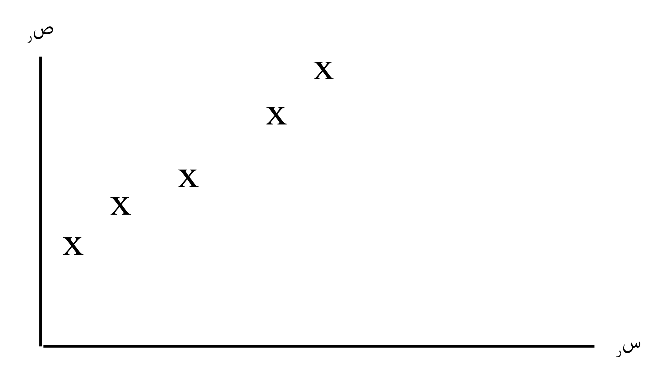
مثال:

البيانات التالي تمثل متغيراً عشوائياً ذا بعـدين هـما، الـوزن بالأرطـال (كغـم) والعمـر بالسـنوات لعينة عشوائية من بعض الصبية المرضى بأحد المستشفيات.

الرقم	الوزن (سر)	العمر (صر)
1	20	4
2	26	7
3	34	11
4	38	13
5	42	5
المجموع	**160**	**50**

يلاحظ من المثال أعلاه أن المتغيرين (سر، صر) يتزايدان في اتجاه واحد، بمعنى أن صر توافق سر في تغيراتها، والرسم البياني الذي يسمى (لوحة الانتشار) يوضح ذلك:

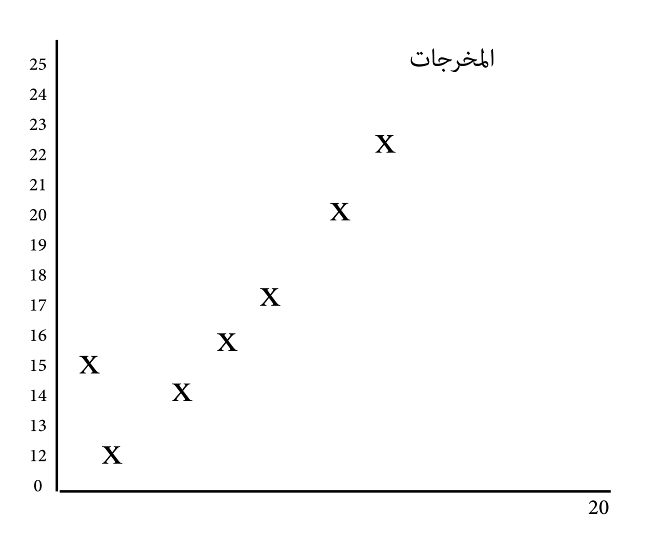
(Scatter Diagram)



البرنامج التالي مثال لكيفية استخدام عبارات بسيطة بلغة بيسك لرسم بياني وهذا المثال يمكن أن يعطي فكرة عن استخدام دوارة For (Loop) وعبارة Print TAB لتحقيق مثل هذا الرسم البياني:

```
10    REM PROGRAM TO POIT VALUES OF A AGAINST X
20    DIM X (7), Y(7), A(50), B(50)
30    READ  N
40    FOR I = 1 TO  N
50    READ X(I) , Y(I)
60    Y(I) = INT (Y0/2 +0.5)
60    Y (I)= INT (Y(I)/2 + 0.5) REM TRANSFORM
70    NEXT  I
80    GO SUB 290   REM TO DETERMINE THE HIGHEST
      & LOWEST VALUES IN Y.
90    GO SUB 370
100   FOR I = H TO L STEP – 1 REM STARTING
      AT HI & ENDING AT LO VALUES OF Y
110   PRINT I. L
```

```
120        IF A(I) >σ THEN PRINT TAB (B(I) + 18); '*'
130        NEXT   I                              ELSE PRINT
140        MAT Y = X
150        GO SUB 290 REM TO DETERMINE THE
           HIGHEST & LOWEST VALUESING X
160        Z = 0 REM DUMMY NUMBER
170        PRINT  Z ;
180        FOR I = 1 TO H
190        PRINT    '-;
200        NEXT  I
210        PRINT
220        PRINT    '     '
230        FOR I ='   1  TO H'
240        PRINT
250        NEXT I
260        PRINT 11
270        DATA  7,10,25,15,35,14,30,16,40,11,24,20,50,19,48
280        stop
290        REM ROUTINE TO DETERMINE HIGHEST
           AND LOWEST VALUES
300        H=Y(I) REM HIGHEST VALUES
310        L = Y(1) REM LOWEST VALUES
320        FOR I = 2 TO N
330        IF Y (I) > L THEN H = Y (I)
340        IF Y(I) < L THEN L= Y (I)
350        NEXT  I
360        RETURN
370        REM ROUTINE TO ARRANGE VALUES OF Y
           IN SEQUENCE IN ANOTHER ARRAYA
380        REM AND ARRANGE COBBESPONDING VALUES
           OF  X  IN ARRAY B
390        FOR I = 1  TO  N
400        A (Y(I)) = Y (I)
410        B (Y(I)) = X(I)
420        NEXT  I
430        RETURN
440        END
```



2/8 التباين المشترك أو (التغاير) COVARIANCE

- هو مجموع مضاريب انحرافات الأزواج العينية المرتبة مقسوما÷ على عددها ناقصاً – واحد.
- يكون التغاير (موجباً) إذا كانت العلاقة طردية، ويكون (سالباً) إذا كانت العلاقة عكسية.
- أما إذا كان التغاير معدوماً (صفراً) فهذا دليل على تساوي مجموع مضاريب الانحرافات الموجبة، بمجموع مضاريب الانحرافات السالبة وهذا يعني عدم وجود علاقة بني المتغيرين.
- التغاير مقياس نوعي وليس كمياً للعلاقة بين المتغيرين.
- التغاير يتغير وحدة القياس، ولا يتأثر بتعديل نقطة الأصل (الجمع والطرح) .
- التغاير الكبير يعني ضعف العلاقة.

$$ع_{س ص} = \frac{\sum س_ر ص_ر - \frac{\sum س_ر \sum ص_ر}{ن}}{ن - 1}$$

مثـال:

استخدم البيانات لإيجاد التغاير ومصفوفة التشتت للمتغيرين $س_ر$،$ص_ر$، والبيانات هي:

$س_ر$	$ص_ر$
20	4
26	7
34	11
38	13
42	15
160	**50**

الحل:

$س_ر$	$ص_ر$	$س_ر^2$	$ص_ر^2$	$س_ر ص_ر$
20	4	400	16	80
26	7	676	49	182
34	11	1156	121	374
38	13	1444	169	494
42	15	1764	225	630
160	**50**	**5440**	**580**	**1760**

$$ع^2_{س} = \frac{\sum س_ر^2 - \frac{(\sum س_ر)^2}{ن}}{ن - 1} = \frac{5440 - \frac{(160)^2}{5}}{4} = 80$$

$$ع^2_{ص} = \frac{580 - \frac{(50)^2}{5}}{4} = 20$$

$$\text{ع س ص} = \frac{\sum \text{س}_\text{ر} \text{ص}_\text{ر} - \frac{\sum \text{س}_\text{ر} \sum \text{ص}_\text{ر}}{\text{ن}}}{\text{ن} - 1} = \frac{1760 - \frac{160 \times 50}{5}}{4} = \frac{160}{4} = 40$$

مصفوفة التشتت هي: $\begin{bmatrix} 40 & 80 \\ 20 & 40 \end{bmatrix}$

البرنامج التالي يقوم بحساب التغاير بين متغيرين، والبيانات المستخدمة في البيانات أعلاه باستخدام معادلة التغاير.

$$C = \frac{S - \frac{X_1 Y_1}{N}}{N - 1}$$

حيث:

$S = \sum X_1 Y_1$

N = حجم العينة

حساب التغاير متغيرين PROGRAM TO CALCULATE THE COVARIANCE

```
10    REM
20    REM
40    X1 = 0
50    Y1=0
60    S=0
70    READ  N
80    FOR I = 1 TO N
90    READ X (I), Y(I)  REM  المتغيرين
100   NEXT  I
110   FOR I = 1 TO N
120   X1 = X1 + X (I)   REM SUM OF X
130   Y1 = Y1 + Y(I)    REM SUM OF Y
140   S = S + X(I) *  Y (I)
150   NEXT   I
```

```
160     C = (S-(X1*Y1) / N) / N-1)
170     PRINT,`  العمر  X  الوزن ,` "العمر ,` " الوزن,` " الرقم,`
180     PRINT ,  ____  ,  _____ ,  ______ , ______ , ______
190     FOR I = 1 TO N
200     PRINT X (I) * Y(I), Y(I) , X(I) , I
210     NEXT  I
220     PRINT, ____ , _________ " ______,
230     PRINT  S,  X1  ,  Y1
240     PRINT , C ,  =  التغاير
250     DATA  5, 20, 4, 26, 7 , 34, 11, 38, 13, 42, 15
260     END
```

المخرجات

الرقم	الوزن	العمر	الوزن × العمر
1	20	4	80
2	26	7	182
3	34	11	374
4	38	13	494
5	42	15	630
	160	**50**	**1760**

3/8 معامل الارتباط الخطي للبيانات النسبيــة

(Pearson's Moment Correlation)

يعرف معامل بيرسون العزمي للارتباط الخطي بأنه القيمة المعيارية للتغـاير، بمعنـى أن معامـل الارتباط الخطي هو: $ر = \frac{ع_{س ص}}{ع_س \; ع_ص}$

$$ر = \frac{\sum س_ر \, ص_ر - \frac{\sum س_ر \sum ص_ر}{ن}}{\sqrt{\left(\sum س_ر^2 - \frac{\left(\sum س_ر\right)^2}{ن}\right)\left(\sum ص_ر^2 - \frac{\left(\sum ص_ر\right)^2}{ن}\right)}}$$

هذا ويلاحظ من المعادلة أعلاه أنه إذا كانت : $س_ر = ص_ر$

فإن: ر = 1

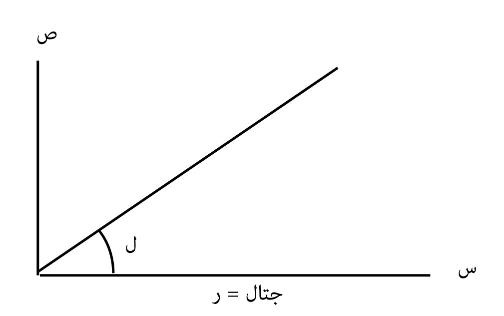
أما إذا كانت : $س_ر = - ص_ر$

فإن: ر = -1

لذلك فللارتباط حدود، غذ أنه يتراوح بين +1 و -1 أي أن:

$$-1 \leq ر \leq +1$$

ويكون الارتباط +1 إذا كانت العلاقة طردية تامة، ويكون -1 إذا كانت العلاقة الخطية عكسية سالبة، ويساوي صفراً إذا كان معدومة، فهو بذلك ذو خصائص مماثلة لخصائص جيب تمام (جتا) الزاوية الواقعة بين الخط المستقيم والمحور السيني (الزاوية ل) في الشكل أدناه .



- لا يتأثر الارتباط بتعديل مقياس الرسم (وحدة القياس)، أو نقطة الأصل (الجمع والطرح).
- ليست له وحدة، قياس، لذلك أصبح معامل الارتباط الخطي هو المقياس الكمي والنوعي للعلاقة الخطية بين المتغيرين، فعدمه أو ضعفه يعني عدم أو ضعف العلاقة الخطية، ولكنه ليس دليلاً على عدم وجود أي علاقة، لأن العلاقة قد تكون غير خطية، كما أن وجوده لا يعني السببية.

•

مثـــال:

استخدم المثال السابق لإيجاد الارتباط بين المتغيرين.

ع س ص = 40 　　 ع^{2}س = 80 　　 ع^{2}ص = 20

$$ر = \frac{ع س ص}{ع س ع ص} = \frac{40}{\sqrt{20 \times 80}} = 1$$

أي أن العلاقة طردية تامة.

البرنامج التالي يقوم بحساب الارتباط الخطي :

$$L = \frac{A}{\sqrt{BC}}$$

حيث:

$$A = S - \frac{X_1 Y_1}{N}$$

$$B = X_2 - \frac{X_1^2}{N}$$

$$C = Y_2 - \frac{Y_1^2}{N}$$

$S = \sum XY$

$X_1 = \sum X$

$Y_1 = \sum Y$

$X_2 = \sum X2$

$Y_2 = \sum U2$

N = حجم العينة

4/8 حساب الارتباط الخطـــي **Program to Compute Linear**

```
10    REM
30    REM
40    X1 = 0
50    Y1 =
60    X2 = 0
70    Y2 = 0
80    S =0
90    READ N REM NO OF OBSERVATIONS
```

```
100        FOR I = 1 TO N
110        READ  X (I),  Y(I)
120        NEXT I
124        FOR I = 1 TO N
125        X1 = X1 + X(I)        REM SUM OF X
130        Y1 = Y1 + Y(I)        REM SUM OF Y
135        X2 = X2 + X(I) *X(I)
138        Y2 = Y2 + Y(I) *Y(I)
140        S = S + X(I) * Y (I)
150        NEXT   I
160        A = S –X1 * Y1 /N
170        B = X2 – X1 * X1 /N
180        C = Y2 – Y1 * Y1/N
190        L = A/SQR (B*C)
195        PRINT  `  صر2 ,`  سر2 ,`  سرصر ,`  صر ,`  سر ,`
198        PRINT ,    _____, _____,  _______,  _____,  _____
200        FOR I = 1  TO  N
210        PRINT Y (I)* Y(I), X(I)* X(I), X(I) , X(I) * Y(I), Y(I) , X(I)
220        NEXT  I
230        PRINT  , ____ ‘;  ______ ‘; ______ ‘; _____ ‘; _____ ‘;
240        PRINT, L.’ =   الارتباط CORRELATION
260        DATA 5, 20, 4,26,7,34,11,38,13,42,15
270        END
```

المخرجات

$ص_{ر}^{2}$	$س_{ر}^{2}$	$س_{ر}\,ص_{ر}$	$ص_{ر}$	$س_{ر}$
16	400	80	4	20
49	676	182	7	26
121	1156	374	11	34
169	1444	494	13	38
225	1764	630	15	42
580	**5440**	**1760**	**50**	**160**

الارتباط = 1

5/8 معنويـــة الارتباط

يعتبر معامل ارتباط بيرسون Pearson مقدراً للارتباط النظري (ز) الخاص بالمجتمع الثنـائي الـذس سبحت منه العينة ذات الحجم (ن)، بيد أن إعادة السحب قد تؤدي إلى معامل ارتبـاط آخـر، وإذا تـم تكرار عملية الاختيار العشوائي للعينات الثنائية، فسوف تكون هناك عدة ارتباطات هـي 3,2,1.... وكـل منها يمثل تقديراً للارتباط الحقيقي الخاص بالمجتمع (ز).

مثال:

اختيرت عينة عشوائية قوامها 27 من بيانات ذات بعدين، فاتضح أن الارتباط الخطي يساوي 0.8 فهل يعتبر ذلك دليلاً على وجود ارتباط بمستوى معنوية 5%؟

الحل:

ف : ز = صفراً

ف1:ز ≠ صفراً

إحصائية الاختبار هي:

$$ت = \frac{ر}{\sqrt{\frac{1 - ر^2}{ن - 2}}} = \frac{0.8}{\sqrt{\frac{1 - 0.64}{25}}} = \frac{5 \times 0.5}{0.6} = 6.67$$

القيمة الحرجة من جدول ت بالملحق (2) في نهاية الكتاب على 25 درجات حرية تساوي 2.060

وبما أن إحصائية الاختبار أكبر من القيمة الحرجة، فلا بد من قبول الفرضـية البديلـة بعـد رفـض فرضية العدم.

مثال:

أختيرت عينة عشوائية حجمها 28 لبيانات ذات بعدين، فاتضح أن الارتباط الخطي يسـاوي 0.23، اختبر الفرضية القائلة بأن الارتباط الحقيقي للمجتمع 0.3، وأوجد حدود الثقة لمعامل ارتباط المجتمـع، وذلك بمستوى معنوية 5%.

الحل:

أ = 5% $\frac{أ}{2} = 2.5\%$

باستخدام جدول التوزيع الطبيعي بالملحق (1) نجد أن القيمة الحرجة هي:

0.975 = 1.96

أما إحصائية الاختيار فهي:

$$ى = \frac{س - و}{م} \quad ، \quad س = \frac{1}{2} \text{ لن } \frac{1 + ر}{1 - ر}$$

ومن جدول (2) في نهاية الفصل يتضح أن:

$$س = \frac{1}{2} \text{ لن } \frac{1 + 0.23}{1 - 0.23} = 0.234$$

وباستخدام نفس الجدول

$$و = \frac{1}{2} \text{ لن } \frac{1 + 0.3}{1 - 0.3} = 0.310$$

$$م = \frac{1}{\sqrt{ن - 3}} = \frac{1}{\sqrt{28 - 3}} = 0.2$$

وعليه تكون إحصائية الاختبار بعد التعويض في المعادلة وهي:

$$ى = \frac{0.234 - 0.310}{0.2} = -0.38$$

وبما أن: $-0.38 > -1.96$

فهذا يعني أنه لا بد من قبول فرضية العدم القائلة بأن ز = 0.3، بمعنى أنه لا يوجد فرق جوهري بين 0.3 والقيمة العينية للارتباط التي تساوي 0.23.

أما حدود الثقة فهي:

$$س \pm ى_{0.975} \times م$$

$$0.234 \pm 1.96 \times 0.2$$

0.392 ± 0.234

وعليه تكون: $- 0.158 \geq , \geq 0.626$

وباستخدام جدول (2) مرة أخرى بطريقة معاكسة لإيجاد ز لكل حالة باعتبار أن:

$$-0.158 = \text{لن} \; \frac{1+\text{ز}}{1-\text{ز}}$$

وأيضاً: $0.626 = \text{لن} \; \frac{1+\text{ز}}{1-\text{ز}}$

يلاحظ أن:

-0.158 تناظرها -0.159

0.626 تناظرها 0.555

وعليه تكون:

$$0.159 \geq \text{ز} \geq 0.555$$

6/8 معامل سبيرمان لارتبـــاط الرتب Spearman's Rank Correlation

1/6/8 معامل ارتباط الرتب لمتغيرين تسلسليين (معامل سبيرمان لارتباط الرتب)

يرتب المتغيران أولاً: تصاعدياً أو تنازلياً، ثم يستخرج الفرق (انحراف كل زوج من الأزواج المرتبة) ويرمز له بالرمز (ل)، وبافتراض أن عدد الفروقات يساوي ن فمعامل سبيرمان لارتباط الرتب هو:

$$\text{ر} = \frac{1 - 6 \sum \text{ل}^2}{\text{ن}(\text{ن}^2 - 1)}$$

هذا، ويمتاز معامل ارتباط سبيرمان للرتب بنفس خصائص معامل بيرسون، وتستخدم نفس الأساليب السابقة لاختبارات الفرضيات وفترات الثقة.

مثال:

البيانات التالية تمثل درجات عينة قوامها 10 أشخاص تقدموا للالتحاق بوظيفة، فأجريت لهم مقابلات شخصية من لجنة مكونة من عنصرين، أوجد كل عضو برصد درجة لكل شخص من 10.

N الرقم	X درجات العضر الأول	Y درجات العضو الثاني
1	9	10
2	5	7
3	6	5
4	1	1
5	3	4
6	2	2
7	8	9
8	7	6
9	10	8
10	4	3

الحل:

N الرقم	X درجات الأول	Y درجات الثاني	D الفرق (ل) الأول - الثاني	D^2 مربع الفرق ($ل^2$)
1	9	10	1-	1
2	5	7	2-	4
3	6	5	1+	1
4	1	1	0	0
5	3	4	1-	1
6	2	2	0	0
7	8	9	1-	1
8	7	6	1+	1
9	10	8	2+	4
10	4	3	1+	1
المجموع			**صفر**	**14**

$$ر = \frac{1-6\sum ل^2}{ن(ن^2-1)} = 1-\frac{6\times 14}{10\times 99} = 0.915$$

البرنامج التالي يقوم بحساب معامل سبيرمان لارتباط الرتب والبيانات المستخدمة بالمثال السابق باستخدام معادلة الارتباط للرتب.

$$L = 1 - \frac{6D^2}{N(N^2 - 1)}$$

حيث:

مجموع مربعات فروقات الرتب = D2

حجم العينة = N

برنامج لحساب معامل سبيرمان لارتباط الرتب

```
10      REM
30      REM  PROGRAM TO COMPUTE SPEARMAN'S BANK ORRELATION
40      D1 = 0
50      D2 = 0
90      READ N  REM NO OF OBSERVATIONS
100     FOR I = 1 TO N
110     READ X (I) , Y(I)
120     D1 = D1 + (X(I) – Y(I))
130     D2 = D2 + (X(I) – Y(I)) **2
140     NEXT I
150     L  = 1 -6 * D2/(N*(N*N-1))
160     PRINT, ل2 , الفرق , التقييم, التقييم , ' رمز,'
170     PRINT  , ___ , ___ , ___, _____, ____,
180     PRINT
190     FOR I = 1 TO N
200     PRINT (X(I) – Y (I)) **2, X(I) – Y(I), Y(I), X(I), I
210     NEXT I
220     PRINT '___, '______,'  '____',  '____', '_____'
230     PRINT  D2,  D1
240     PRINT   L  ' = CORRELATION COEFFICIENCE  معامل الارتباط
250     DATA 10,9,10,5,1,7,6,5,1,1,3,4,2,2,8,9,7,6,10,8,4,3
300     END
```

المخرجات

رمز المتقدم	التقييم الأول	التقييم الثاني	الفرق ل	$ل^2$
1	9	10	1-	1
2	5	7	2-	4
3	6	5	1+	1
4	1	1	0	0
5	3	4	1-	1
6	2	2	0	0
7	8	9	1-	1
8	7	6	1+	1
9	10	8	2+	4
10	4	3	1+	1
المجموع			**صفر**	**14**

معامل الارتباط = 0.9151515

2/6/8 الارتباط الثنائي التسلسل Bisearial Correlation

إذا كان المتغير $س_ر$ متصلاً، بينما كان المتغير $ص_ر$ ثنائي التسلسل، كأن تقسم المدينة إلى منطقتين، أو حالات الإجابة بنعم أو لا، أو الحالة الاجتماعية (متزوج، أو غير متزوج، وإذا كانت:

$\overline{س}_1$ = الوسط الحسابي للمتغير المتصل عندما كانت $ص_ر = ص_1$.

$\overline{س}_2$ = الوسط الحسابي المقابل لصفة التقسيم الثانية ($ص_2$)

$$ح = \frac{ص_1}{ص_1+ص_2} \quad ، \quad \overline{ح} = \frac{ص_2}{ص_1+ص_2} = 1-ح$$

فالارتباط الخطي بني المتغير النسبي المتصل ($س_ر$) والمتغير الخاص بصفتي التقسيم ($ص_ر$) هو:

$$ر = \left(\frac{س_1-س_2}{ع\ س}\right)\times\left(ح\overline{ح}\right)^{\frac{1}{2}}$$

ويلاحظ أن المتغير الثنائي التسلسل لا يدخل في العمليات الحسابية الخاصة باستخراج الارتباط.

أما اختبارات المعنوية وحدود الثقة، فتستخدم لها نفس المعادلات الخاصة بمعامل بيرسون للارتباط الخطي.

مثال:

البيانات التالية تمثل عدد القتلى في حوادث المرور، موزعين حسب مكان الوفاة خلال الفترة من إبريل حتى ديسمبر 1980، والبيانات هي:

الشهر	عدد القتلى قبل الوصول للمستشفى	عدد القتلى بعد الوصول للمستشفى
إبريل	21	18
مي	26	21
جون	36	26
جلاي	46	33
أغسطس	39	29
سبتمبر	34	24
أوكتوبر	30	19
نوفمبر	19	14
ديسمبر	19	15
	270	**199**

أوجد الارتباط بين عدد القتلى ومكان الوفاة.

الحل:

$$\text{ر} = \frac{\bar{\text{س}}_1 - \bar{\text{س}}_2}{\text{ع س}} \times \sqrt{\text{ح}\,\bar{\text{ح}}}$$

$$\text{ع س} = 8.868$$

$$\bar{\text{س}}_1 = \frac{270}{9} = 30 \quad ، \quad \bar{\text{س}}_2 = \frac{199}{9} = 22.1$$

$$\text{ح} = \frac{\text{ص}_1}{\text{ص}_1 + \text{ص}_2} = \frac{270}{199 + 270} = 0.576$$

$$\bar{\text{ح}} = \frac{\text{ص}_2}{\text{ص}_1 + \text{ص}_2} = \frac{199}{469} = 0.424$$

$$\therefore \text{ر} = \frac{30 - 22.11}{8.868} \times \sqrt{0.424 \times 0.576} = 0.44$$

يلاحظ من جدول (1) في نهاية هذا الفصل أن القيمة الحرجة لاختبار الفرضية:

ف$_0$:صفراً مع الفرضية البديلة ف$_1$:ز> صفر

وبمستوى 5% تساوي 0.497، وفي ذل دلالة على خطورة الإصابات التي يصعب إسعافها في كثير من الحالات.

البرنامج التالي يقوم بحساب الارتباط الثنائي مستخدماً البيانات أعلاه، وباستخدام المعادلة:

$$D = \frac{B - C}{V} \sqrt{AE}$$

حيث:

$$B = \frac{X_1}{N}$$

$$C = \frac{Y_1}{N}$$

$$A = \frac{X_1}{X_1 + Y_1}$$

$$V = \sqrt{\left(X_2 + Y_2 - (X_1 + Y_1)/N\right)/(N-1)}$$

$X_1 \; Y_1 =$ مجاميع

$X_2 \; Y_2 =$ مجاميع مربعات

لا بد من حساب الانحراف المعياري (V) هنا الذي يساوي 8.868 المستخدم في السطر رقم 220 وهو عبارة عن

$$\begin{bmatrix} (21)^2 + (26)^2 + (36)^2 + ... + (19)^2 + (18)^2 + (21)^2 + (15)^2 \\ - \sum (21 + 26 + 36 + ... + 15)^2 \;\; / \; N \end{bmatrix}$$

حيث N= 18 ÷N-1

برنامج لحساب الارتباط الثنائي:

```
10     REM
30     REM PROGRAM TO COMPUTE BISERIAL CORRELATION
40     X1 = 0
50     Y2 = 0
60     READ N  REM NO OF OBSERVATIONS
70     PRINT , العدد2 , العدد1 ,  الشهر '
80     PRINT , ___ , _____, ____,
90     FOR I = 1  TO  N
100    READ  X,Y
110    PRINT  Y, X, I
120    X1 = X1 + X
130    Y1 = Y1 + Y
140    NEXT  I
150    A = X1 / (X1+Y1)
190    B = X1/N
200    C= Y1/N
210    E = 1-A
220    D = (B-C)/ 8.868* SQR (A*E)
230    PRINT  ,________,' , '_____, '   '
240    PRINT  Y1 , X1,' المجموع  TOTAL  SUMMATION
250    PRINT, D.,' =   الارتباط الثنائي  BISERIAL CORRELATION
260    DATA 9,12,18,26,21,36,26,46,33,39,29,34,24,30,19,19,14,19,15
300    END
```

المخرجات

الشهر	العدد 1	العدد 2
1	21	18
2	26	21
3	36	26
4	46	33
5	39	29
6	34	24
7	30	19
8	19	14
9	19	15
المجموع	**270**	**199**

الارتباط الثنائي Biserial Correlation = 0.4396694

الوحدة التاسعة

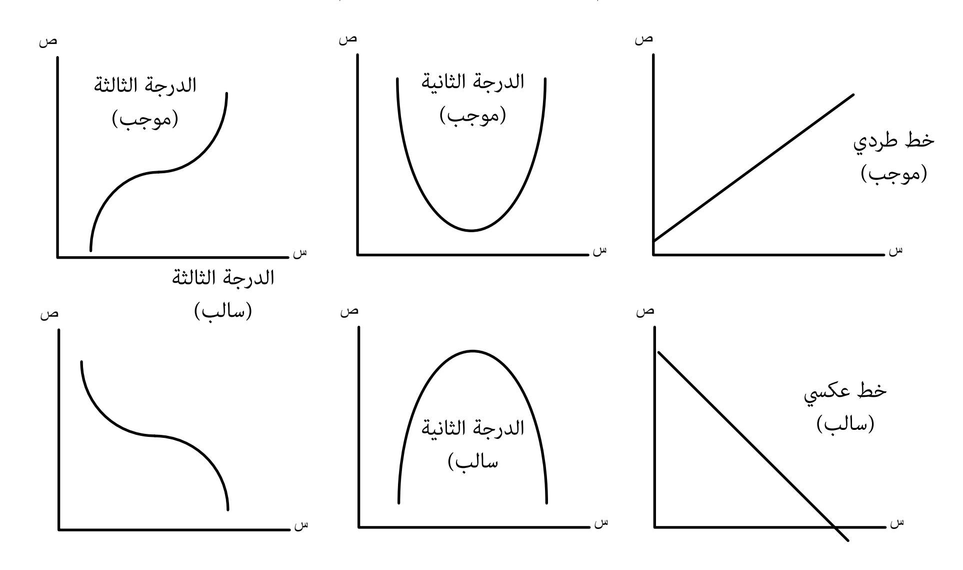
الانحدار الخطـــــــي
(*Linear Regression*)

1- مفهوم الانحدار

يهدف خط الانحدار الخاص بمتغيرين متصلين إلى تحديد العلاقة بين القيم العينية الثنائية (سر،صر) باعتبار أن سر دالة للمتغير صر، بمعنى أن المتغير سر مسبب (مستقل). والتغير صر متغير تابع Dependent Variable

يتلخص دور الانحدار في ثلاث مهام أساسية.

1. الوصف الخاص بظاهرة معينة.
2. التحكم في المتغير التابع بواسطة المتغير المستقل.
3. تقدير (التنبؤ) بعض قيم المتغير التابع بعد تحديد قيم معينة.



نماذج لبعض المنحنيات

ويكون الانحدار الخطي على النحو التالي:

$ص_ر$ = (ا) + (ب) $س_{1ر}$ + (جـ) $س_{2ر}$ + ـ5) $س_{3ر}$ + ...خ.

حيث: ص ر متغير تابع $س_1$، $س_2$، $س_3$ ... مجموعة المتغيرات المستقلة.

خ = الخطأ العشوائي (أ)ن (ب)، (جـ) ... هي معالم (ثوابت) المعادلة الواجب تقديرها.

معادلة الانحدار الخطي البسيط:

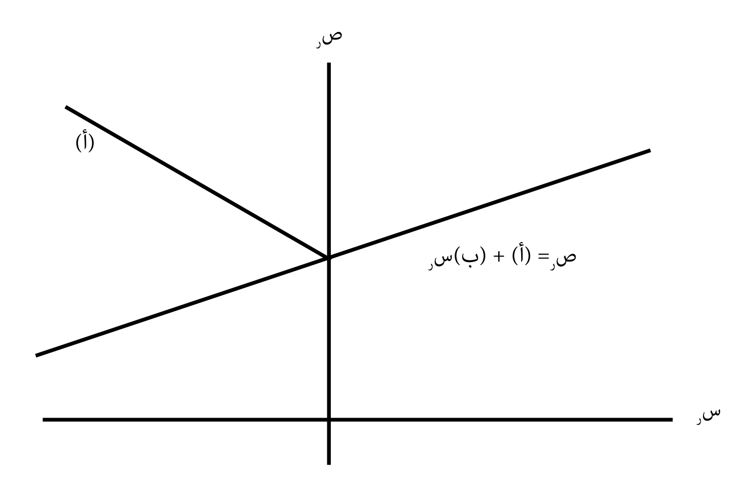
تكون معادلة الانحدار الخطي البسيط على النحو التالي:

$ص_ر$= (أ) + (ب) $س_ر$

حيث:

(أ) هي الجزء المقطوع من المحور الصادي، أو هي المعدل العام.

(ب) هي ميل خط الانحدار، أو كمية التغير التي تطرأ على المتغير التابع ($ص_ر$) إذا تغير المستقل ($س_ر$) بوحدة واحدة.



خطة الانحدار البسيط

الانحدار: هو تمثيل للعلاقة المتوسط بين المتغيرات، فإذا وجدت علاقة بين المتغيرات المطلوب دراستها، يمكن توفيق معادلة لمنحنى أو لخط يحدد طبيعة تلك العلاقة. ومن ثم يمكننا استخدام المعادلة للتنبؤ، أي لتقدير القيم النظرية لمتغير (المتغير التابع) والتي تقابل قيم معينة للمتغيرات المستقلة.

ومدى اقتراب القيم النظرية والحقيقية، يعتبر مقياساً لقوة تمثيل المنحنى أو الخط الموافق للعلاقة بين المتغيرات.

وتسمى المنحنيات أو المعادلات التي تربط بين المتغيرات بمنحنيات، أو معادلات الانحدار.

أبسط صور الانحدار هو: الاندار الخطي أو الانحدار المستقيم، والذي يصف علاقة من الدرجة الأولى بين المتغيرين تحت الدراسة.

ولتوضيح العلاقة بين الانحدار والارتباط، نجد أن هناك عدة طرق:

1/9 الرسم البيانـــي: Graphic

يمكن عرض قيم الظاهرتين برسم شكل الانتشار، ونمهد باليد الخط المستقيم الذي يمر بأكبر عدد من النقاط، هذا الخط يصف العلاقة المتوسطة بين المتغيرين.

وبالطبع ليس ضرورياً أن تقع كل النقاط المصورة لأزواج القيم على الخط.

إذا كان الارتباط قوياً بين المتغيرين، ستقع أغلب النقاط على الخط أو قريبة منه.

فإذا اعتبرنا أن Y هو المتغير التابع Dependent و x هو المتغير المستقل Independent Variable فإن معادلة خط الانحدار يمكن تمثيلها بالمعادلة:

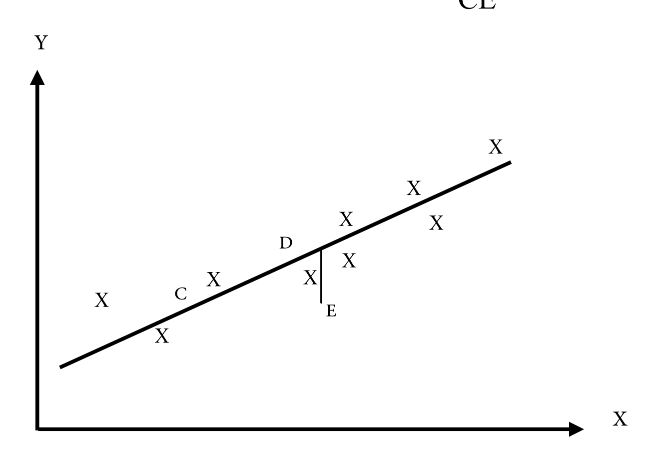
$$\hat{Y} = A + BX$$

حيث A, B ثوابت.

إذا أمكننا تقدير قيم الثوابت أو المعالم Parameters من الرسم أمكننا استنباط المعادلة المطلوبة، وهذه تسمى معادلة انحدار $\frac{X}{Y}$ حيث جرت العادة على وصف الخط باسم المتغير التابع أوّلاً.

(A) يمثل طول الجزء الذي يقطعه الخط الممهد من المحور الرأسي، ويمكن استنتاجه مباشرة من الرسم، حيث يوضح لنا قيمة Y عندما تكون X.

(B) تمثل معدل التغير في المتغير التابع. حيث يتغير المتغير المستقل بوحدة واحدة، وهي ميل الخط، ويساوي = ظل الزاوية التي يصنعها الخط مع المحور الأفقي، وهذه يمكن حسابها بأخذ نقطتين على الخط مثل C,D في الرسم، ونرسم من النقطة D خطاً موازياً // للمحول الرأسي، ومن النقطة C خطاً موازياً// للمحور الأفقي، حيث يلتقوا في النقطة E، ويكون ظل هذه الزاوية مساوياً لقسمة ÷ المقابل على المحاور $\frac{DE}{CE}$.



وبالحصول على قيمتي المعالم، تتحدد معادلة الخط الذي يصف العلاقة بين الظاهرتين.

يلاحظ بأن الخط الممهد بهذه ا لطريقة تحكى، وتعوزه الدقة إلى حد بعيد ويختلف من باحث إلى آخر حسب خبرته ومرانه، لذلك لا يمكن الاعتماد عليها.

2/9 طريقة المربعات الصغرى Ordinary Least Squares

يمكن تقدير قيمة المعالم A,B في خـط الانحـدار $\hat{Y} = A + BX$ باسـتخدام طريقـة المربعـات الصغرى تعطينا أفضل توفيـق للخـط المتوسـط Best Fit، وفيهـا يتحـدد خـط الانحـدار، بحيـث يكـون مجموع مربعات انحرافات القيم المشاهدة (الفعلية) عنه أصغر ما يمكن، أي أن هـذا المجمـوع يكـون أقل من مجموع مربعات انحرافات هذه المشاهدات عن أي خط آخر.

كما يتحدد خط الانحدار أيضاً عندما تكون مجموع انحرافات القيم عن الخط = تساوي صـفر، أي أن الانحرافات موجبة، أي المشاهدات التي تقع فوق الخط، تواجهها انحرافات سالبة تقع أسفل الخط، ويمكننا التعبير عن هذين الشرطين بأن:

$$\sum\left(Y - \hat{Y}\right) = \text{صفر}$$

$$\sum\left(Y - \hat{Y}\right)^2 = \text{أقل ما يمكن}$$

حيث $\hat{Y}$ تمثل القيم المقدرة باستخدام دالة الانحدار.

Y القيم الفعلية للمتغير.

وبالتعويض عن Y بمعادلة الانحدار، في الشرط الثاني، واستخدام التفاضـل الجـزئي، يمكـن إيضـاح أنه لكي يكون الشرط الثاني نهاية صغرى، يلزم أن يكون:

$$\sum Y = NA + B\sum X$$

$$\sum XY = A\sum X + B\sum X^2$$

حيث أن:

N = عدد القراءات بالعينة.

$\sum Y$ = مجموع قيم المتغير التابع.

$\sum X$ = مجموع قيم المتغير المستقل.

A = الجزء المقطوع من المحور الرأسي.

B = ميل خط الانحدار

$\sum XY$ = مجموع حواصل ضرب أزواج القيم المتناظرة للمتغيرين.

وبحل المعادلتين انيًا على القانونين اللذان يعطينا قيم الثوابت:

$$B = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - \left(\sum X\right)^2} \quad , \quad A = \overline{Y} - B\overline{X}$$

وبالمثل يمكننا تقدير معادلة انحدار $\frac{Y}{X}$، أي أن X ستصبح المتغير التابع، وأن Y سيكون التغير المستقل.

ونود تقدير قيمة X المتوسطة المقابلة لقيمة معينة للمتغير Y.

وتكون المعادلة المطلوب تقدير معالمها، هي:

$$\hat{X} = H + WY$$

ولتقدير قيم الثوابت H,W نحل المعادلتين

$$\sum X = NH + W\sum Y \qquad (1)$$

$$\sum YX = H\sum Y + W\sum Y^2 \qquad (2)$$

أو يمكن التعويض في القوانين التالية، وهي ناتجة عن حل المعادلتين السابقتين

$$B = \frac{N\sum XY - \sum X \sum Y}{N\sum X^2 - \left(\sum X\right)^2} \quad , \quad H = \overline{X} - W\overline{Y}$$

ويمكن توضيح كيفية إيجاد معادلتي الانحدار باستخدام طريقة المربعات الصغرى، باستخدام المثال التالي:

مثال:

احسب معادلتي الانحدار $\frac{Y}{X}, \frac{X}{Y}$ من البيانات التالية:

X: 12 0.8 1.0 1.3 0.7 0.8. 1.0 0.6 0.9 1.1

Y:101 92 110 120 90 82 93 75 91 105

الحل:

X	Y	XY	X^2	Y^2
1.2	101	121.2	1.44	10201
0.8	92	73.6	0.64	8464
1.0	110	110.0	1.00	12100
1.3	120	156.0	1.69	14400
0.7	90	63.0	0.44	68100
0.8	82	65.0	0.64	6724
1.0	93	93.0	1.00	8644
0.6	75	4.50	0.36	5625
0.9	91	81.9	0.11	8281
1.1	105	115.1	1.21	11025
9.4	**959**	**924.4**	**9.28**	**93569**

أولاً: معادلة انحدار $\frac{X}{Y}$

(أ) المعادلتين المطلوب حلهم آنياً، هم:

$$\sum Y = AN + B\sum X$$

$$\sum XY = A\sum X + B\sum X^2$$

959 = 10A + 9.4B

924.4 = 9.4A + 9.28B

بحل المعادلتين، نجد أن A = 47.33 , B = 51.67 :

وتكون المعادلة هي:

$$\therefore \hat{Y} = 47.33 + 51.67X$$

(ب) يمكن التعويض في القوانين مباشرة:

$$B = \frac{N\sum XY - \left(\sum X\right)\left(\sum Y\right)}{N\sum X^2 - \left(\sum X\right)^2}$$

$$A = \overline{Y} - B\overline{X}$$

$$B = \frac{(10)\ (92.4.4)\ -\ (9.4)\,(959)}{(10)(9.28) - (9.4)^2}$$

$= 51.67$

$A = 95.9 - (51.67)\ (0.94) = 47.33$

$\hat{Y} = 47.33 + 51.67X$

ثانياً: معادلة انحدار $\frac{X}{Y}$ **، وهي** $\hat{X} = H + WY$

يمكن حسابها بالتعويض في المعادلتين:

$\sum X \quad = NH + W\sum Y$

$\sum YX = H\sum Y + W\sum Y^2$

وباستخدام الجدول السابق نجد أن:

$9.4 \quad = 10H + 959W$

$924.4 = 959H + 93569\ W$

وبحل المعادلتين نجد أن:

$H = 0.403 \quad W = 0.01$

أي أن المعادلة المطلوبة هي : $\hat{X} = -\,0.403 + 0.014\ Y$

ويمكن حسابها بالتطبيق في القانون مباشرة:

$$W = \frac{N\sum XY - \left(\sum X\right)\left(\sum Y\right)}{N\sum Y^2 - \left(\sum Y\right)^2}$$

$H = \overline{X} - W\overline{Y}$

$$W = \frac{(10)\,(924.4)\,(9.4)\,(959)}{(10)(93569) - (959)^2} = 0.014$$

$H = 90.4 - (0.014)\ (95.9)$

$= -0.403$

$\hat{X} = -0.403 + 0.014\ Y$

عموماً يختلف معادلتي الانحدار ولا يتطابقا إلا إذا الارتباط بين المتغيرين تام.

إذا علمت أنه في دراسة عينة مكونة من (12) مشاهدة كان:

$$\sum X = 725 \,,\ \sum Y = 1011,\ \sum XY = 61685,\ \sum X^2 = 44475.$$

فاحسب معادلة انحدار $\frac{Y}{X}$

الحل:

نعوض في القانون للحصول على قيم A,B في المعادلة $\hat{Y} = A + BX$

$$B = \frac{(12)(61685) - (725)(1011)}{(12)(44475) - (725)^2}$$

$$= 0.897$$

$$A = \left(\frac{1011}{12}\right) - (0.897)\left(\frac{725}{12}\right)$$

$$= 30.056$$

وبالتالي يكون خط الانحدار المطلوب، هو:

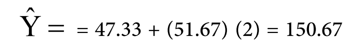
$$\hat{Y} = 30.056 + 0.897\ X$$

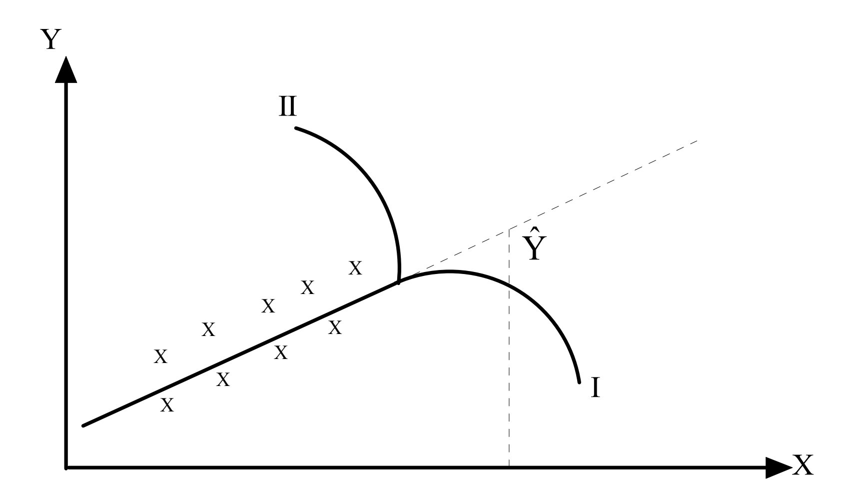
يمكن استخدام المعادلة التي توصلنا إليها للتنبؤ بقيمة المتغير التابع عندما يأخذ المتغير المستقل قيمة محددة التي قد تكون من ضمن القراءات الموجودة بالعينة أو من خارج حدود العينة.

بالرجوع على معادلة الانحدار $\frac{Y}{X}$، من المثال الأول، أي:

$$\hat{Y} = \ = 47.33 + 51.67X$$

كما يمكن استخدام المعادلة في حساب القيمة التقديرية للمتغير التابع عندما يأخذ المتغير المستقل قيمة تقع خارج حدود العينة. باستخدام نفس المعادلة السابقة، فإنه عنده ج= X فإن :


$\hat{Y} = 47.33 + (51.67)\ (2) = 150.67$



مثـــال:

البيانات التالية تمثل أرباح إحدى المؤسسات خلال (7) سنوات، وتكلفة الدعاية في كل سـنة مـن تلك السنوات، أوجد معادلة الانحدار ا لخطي البسيط للربح على الدعاية، ثـم اسـتخدم تلـك المعادلـة لتقدير الربح إذا كانت تكلفة الدعاية 18 ألف دينار في سنة ما، والبيانات هي:

الحل:

الدعاية ($س_ر$) آلاف الدنانير	الربح ($ص_ر$) بآلاف الدنانير
10	25
15	35
14	30
16	40
11	24
20	50
19	48

يتضح من البيانات السابقة أن: ن =7

أما بقية المجاميع فيمكن الحصول عليها على النحو الآتي:

رقم المشاهدة	$س_ر$	$ص_ر$	$س_ر^2$	$ص_ر^2$	$س_رص_ر$
1	10	25	100	625	250
2	15	35	225	1225	525
3	14	30	196	900	420
4	16	40	256	1600	640
5	11	24	121	0.576	264
6	20	50	400	2500	1000
7	19	48	361	2304	912
المجموع	**105**	**252**	**1659**	**9730**	**4011**

$$ع_{س ص} = \frac{\sum س_ر ص_ر - \frac{\sum س_ر \sum ص_ر}{ن}}{ن-1} = 38.5$$

$$ع^2_س = \frac{\sum س_ر^2 - \frac{(\sum س_ر)^2}{ن}}{ن-1} = 14$$

$\therefore$ ع س = 3.742

$\bar{س} = 15$

$\bar{ص} = 36$

$$ب = \frac{ع_{س ص}}{ع^2_س} = \frac{38.5}{14} = 2.75$$

$أ = \bar{ص} - ب\bar{س} = 36 - 2.75 \times 15 = 5.25$

وبالتعويض في المعادلة السابقة $\hat{ص}_ر = أ + ب س_ر$

$\hat{ص}_ر = 5.25 + 2.75 س_ر$

تقدير الربح إذا بلغت تكلفة الدعاية 18 ألف دينار هو:

$\hat{ص}_ر = -5.25 + 2.75 \times 18 = 44.250$ ألف دينار

= 44250 دينار

البرنامج التالي يقوم بإيجاد معادلـة الانحـدار الخطـي البسـيط، البيانـات المسـتخدمة بالبرنـامج والبيانات الواردة في المثال السابق، باستخدام المعادلة :

$$Y = A + BX$$

وباعتبار أن:

الوسط الحسابي للمتغير المستقل = X_3

الوسط الحسابي للمتغير التابع = Y_3

مجموع مربعات المتغير المستقل = X_2

مجموع المتغير المستقل = $_x1$

وبالتالي:

$$B = E/F$$

التغاير $E = (S - X_1Y_1/N)\ (N-1)$

تباين المتغير المستقل = F

$A = Y_3 - BX_3$

برنامج إيجاد معادلة الانحدار الخطي البسيط:

```
10      REM      PROGRAM OF LINEAR REGRESSION
20      DIM X(7) , Y(7)
30      X1 = 0
40      Y1 = 0
50      X2 = 0
60      Y2 = 0
70      READ  N,V
80      FOR I = 1   TO  N
90      READ X (I) , Y(I)
100     X1 = X1 + (I)
110     Y1 = Y1 + Y(I)
120     X2 = X2 + X(I) * X(I)
130     Y2 = Y2 +Y(I) * Y(I)
```

```
140     S = S + X(I) * Y(I)
150     NEXT   I
160     E = (S-X1 * Y1/N) / (N-1)
170     F = (X2-X1 * X1/N)/ (N-1)
180     F1 = SQR (F)
190     X3 = X1/N  REM MEAN OF X
200     Y3 = Y1/N  REM MEAN OF Y
210     B = E/F
220     A = Y3 – B*X3
230     R = A+B*V
240     REM PRINTING SECTION
245     PRINT USING 330
250     PRINT USING 320
260     PRINT USING 330
270     FOR I = 1 TO N
280     PRINT USING 340,X(I)*Y(I),Y(I)* Y(I),X(I) * X(I), Y(I), X(I),1
290     NEXT   I
300     PRINT USING 330
310     PRINT USING  350, S, Y2, X2, Y1, X1
315     PRINT USING 330
320     :          رقم     سر     صر     س2ر     ص2ر     سر صر
330     :          ________________________________________
340     :          *****     ****     ****     **   **     **
350     :          *****    *****     *****    ***     ***  المجموع
360     PRINT
370     PRINT , A`,  ‘= أ ‘
380     PRINT, B;  ‘  =  ب ‘
390     PRINT
460     DATA 7, 18,10,25,15,35,14,30,16,40,11,24,20,50,19,48
500     END
```

المخرجات

رقم	س$_{ر}$	ص$_{ر}$	س$_{ر}^{2}$	ص$_{ر}^{2}$	س$_{ر}$ص$_{ر}$
1	10	25	100	625	250
2	15	35	225	1225	525
3	14	30	196	900	420
4	16	40	256	1600	640
5	11	24	121	0.576	264
6	20	50	400	2500	1000
7	19	48	361	2304	912
المجموع	**105**	**252**	**1659**	**9730**	**4011**

أ= -5.25

ب= 2.75

خصائص معادلة الانحدار الخطي البسيط

أ- المعدل العام

ص ر = أ + ب س$_{ر}$

أ = $\overline{ص}$ - ب $\overline{س}$ بتعويض المعادلة الثانية في الأولى

$\overline{ص}$= $\overline{ص}$ - ب س+ ب س$_{ر}$

$\overline{ص}$ = $\overline{ص}$ + ب (س$_{ر}$- $\overline{س}$)

فإذا كانت س$_{1}$ = س$_{2}$ س$_{3}$.

فإن $\overline{ص}_{ر}$=$\overline{ص}$ أي أن $\overline{ص}$، هي المعدل العام (الوسط الحسابي) إذا كانت س$_{ر}$ ليست متغيرة.

ب- خط الانحدار يمر بالنقطة ($\overline{س}$،$\overline{ص}$).

ج- مربع الارتباط الخطي هو المقياس لدقة التقدير.

3/9 معامل التحديد Coefficient of Determination

يسمّى مربع معامل الارتباط R_2 بمعامل التحديد الذي يعرف بأنه نسبة التغير المفسر إلى التغـير الكلي في أحد الظواهر أو المتغيرات، أي أن:

$$R_2 = \frac{\text{التغير المفسر}}{\text{التغير الكلي}}$$

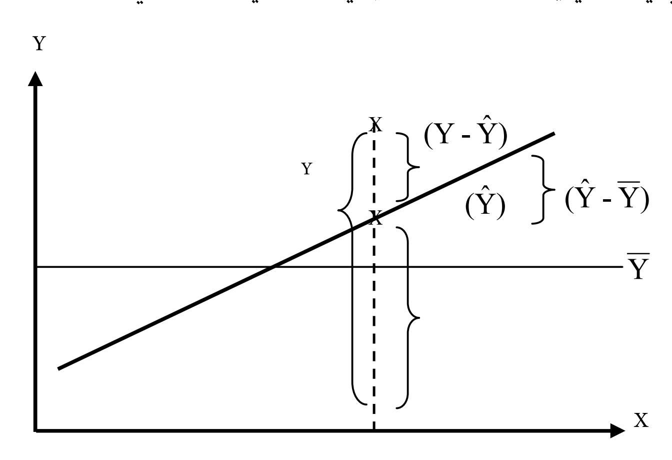
ويعـرف التغـير الكـلي لأي ظـاهرة بأنـه: مجمـوع مربعـات الانحرافـات عـن الوسـط الحسـابي للظاهرة.

فالتغير الكلي في Y مثلاً يساوي = $\left(Y - \overline{Y}\right)^2$، وهذا التغير الكلي يمكن التعبير عنه بالمعادلة:

$$\left(Y - \overline{Y}\right)^2 = \left(Y - \hat{Y}\right)^2 + \left(\hat{Y} - \overline{Y}\right)^2$$

وتسمى القيمة $\left(Y - \hat{Y}\right)^2$ بالتغير الغير مفسر، أي الذي لا يعود إلى التغير في المتغير المستقل.

بينما تسمى القيمة $\left(\overline{Y} - \hat{Y}\right)^2$ بالتغير المفسر بالتغيرات في المتغير المستقل، أي أنه الجزء مـن التغير الكلي في Y الذي يمكن إرجاعه إلى التغير في X ، كما في الشكل التالي:



- فالجزء الأول $\left(Y - \hat{Y}\right)^2$ يمثل مجموع مربعات المسافات الرأسية التـي تمثـل انحرافـات القـيم المشاهدة عن خط الانحدار.
  إن التغير في المتغير المستقل لا يفسر هذا الجزء من التغير الكلي.
- في حين أن الجزء الثاني $\left(\hat{Y} - Y\right)^2$ يقيس الجزء من التغير الكلي الذي يعود إلى المتغير المستقل.
- إن التغير (X) المتغير المستقل يشرح كل التغير في Y، ويكون معامل التحديد مساوياً الواحد. "أي أن معامل التحديد يوضح نسبة التغـير في المتغـير التـابع التـي يمكـن إرجاعهـا للتغـير في المتغـير المستقل".

وبالتالي، فإنه يمكن كتابة معامل التحديد بالشكل التالي:

$$R^2 = \frac{\sum(\hat{Y} - \overline{Y})^2}{\sum(Y - \overline{Y})^2}$$

كما يمكن إعادة كتابتها بالتعويض عن البسط، بقيمتها من المعادلة:

$$R^2 = \frac{\sum(\hat{Y} - \overline{Y})^2 - \sum(Y - \hat{Y})^2}{\sum(Y - \overline{Y})^2}$$

$$= 1 - \frac{\sum(Y - \hat{Y})^2}{\sum(Y - \overline{Y})^2}$$

$$= 1 - \frac{\sigma\hat{Y}}{\sigma Y}$$

حيث R^2 معامل التحديد

$\sigma\hat{Y}$ تباين القيم المقدرة

σY تباين المتغير Y.

وكذلك يمكن حساب معامل التحديد (وحساب معامل الارتباط بأخـذ الجـزر التربيعـي) إذا مـا استخدمنا معادلة انحدار $\frac{Y}{X}$ وحسبنا تباين القيم المقدرة، وتباين القيم الفعلية من المعادلة:

$$R^2 = 1 - \frac{\sigma\hat{X}}{\sigma X}$$

مثال:

احسب معامل التحديد باستخدام معادلة انحدار $\frac{X}{Y}$ من المثال السابق؟

$$\sigma X = \frac{\sum X^2 - H\sum X - W\sum XY}{N}$$

$$= \frac{9.28 - (-0.403)(4.4) - (0.014)(924.4)}{10}$$

$$= 0.0.13$$

$$\sigma X = \frac{\sum X^2}{n} - \left(\frac{\sum X}{N}\right)^2$$

$$= 0.928 - (0.94)^2$$

$$R^2 = 1 - \frac{0.013}{0.045} = 0.72$$

1/3/9 تطبيقات معامل التحديد على الحاسوب

معامل التحديد **Coefficient of Determination**

يطلق على النسبة المئوية التي تفسرها معادلة الانحدار بمعامل التحديـد، وهـي قيـاس لمسـتوى دقة المعادلة، لأنها تمثل عدد النقاط الواقعة على الخط من كل مائة نقطة، وبذلك يمكن تفسير معامل التحديد بأنه = $\frac{\text{م م ر}}{\text{م م ك}}$

أي أن معامل التحديد =

$$\frac{\sum(\text{ص}_{\text{ر}} - \overline{\text{ص}})^2}{\sum(\text{ص}_{\text{ر}} - \overline{\text{ص}})^2} \times 100$$

$$\text{معامل التحديد} = \frac{(\text{ع س ص})^2}{\text{ع}^2\text{س} \times \text{ع}^2\text{ص}} \times 100 = \frac{\text{ع س ص}}{\text{ع س ع ص}} \times \frac{\text{ع1س ص}}{\text{ع س ع ص}} \times 100$$

معامل التحديد = ر × ر × 100= 100 $ر^2$.

= 100 × مربع الارتباط الخطي

مثال:

أوجد معامل التحديد لخط النحدار الوارد في المثال السابق.

الحل:

ع س = 3.742

ع ص = 10.472

ع س ص 38.5

$$\text{ر} = \frac{\text{ع س ص}}{\text{ع س ع ص}} = \frac{38.5}{3.742 \times 10.472} = 0.982$$

معامل التحديد = 100 $ر^2$

= 100 × $(0.982)^2$ = 96.4 %

بمعنى أن حوالي 96% من كل مائة نقطة تكون على خط الانحدار.

البرنامج التالي يقوم بحساب "معامل التحديد" لخط الانحدار السابق اعتماداً على أن معامل التحديد هو مبع الارتباط الخطي.

حيث الارتباط الخطي هو:

$$R = \frac{E}{FG}$$

الانحراف المعياري للتغير المستقل = F

الانحراف المعياري للمتغير التابع = G

برنامج لإيجاد معامل التحديد:

```
10      REM    COEFFICIENT  OF DETERMINATION
20      N = 7
30      S = 4011
40      X1 = 105      REM  SUM   OF X
50      X2 = 1659    REM  SUM  OF  X    SQUARE
60      Y1 = 252      REM  SUM  OF  Y
70      Y2 = 9730    REM SUM OF Y SQUARE
80      E = (S-X11*Y1/N) / (N-1)
90       F = SQR ((X2-X1 *X1/N)/(N-1))
100     G = SQR ((Y2-Y1*Y1/N)/(N-1))
110     R = E7 (F*G)
120     PRINT, E; ‘= ع س ص
130     PRINT
140     PRINT,  F ;  ‘ =  ع س
150     PRINT
160     PRINT , G; ` = ع ص
170     PRINT
180     PRINT, R * R; ` =  إذن معامل التحديد
190     PRINT
200     END
```

المخرجات

ع س ص = 38.5

ع س = 3.741657

ع ص = 10.47218

إذن معامل التحديد = 0.965426

انحرافات التقديرات

الخطأ المعياري لخط الانحدار (ع):

تسمى الإحصائية $\frac{\text{م م خ}}{\text{ن - 2}}$ بتباين خط الانحدار ($ع^2$) أو تباين الخطأ RESIDUAL أي أن:

$$ع^2 = \frac{\sum(ص_ر - \bar{ص}_ر)^2}{ن - 2}$$

بينما يسمى الجذر التربيعي موجب لتباين الخطأ بالخطأ المعياري للتقدير، إذا فالخطأ المعياري لتقدير ص هو

$$ع = \sqrt{\frac{\sum(ص_ر - \bar{ص}_ر)^2}{ن - 2}}$$

$$= \sqrt{\frac{\sum(ص_ر - أ \sum ص_ر - ب \sum س_ر ص_ر)^2}{ن - 2}}$$

الانحدار الثنائي:

الانحدار الثنائي هو أحد أنواع الانحدار المتعدد فنموذجه يتكون من متغيرين مستقلين فقط، أي

$ص_ر = أ + ب\, س_{1ر} + جـ\, س_{2ر}$

البرنامج التالي يقوم بإيجاد معادلة الانحدار الخطي البسيط، وجدول تحليل التبـاين باسـتخدام المصفوفات.

التعليمات الخاصة بالمصفوفات مثل: MAT READ

TRN

INV "معكوس المصفوفة"

برنامج لإيجاد معادلة الانحدار الخطي البسيط وجدول تحليل التباين باستخدام المصفوفات

```
10    REM
15    REM
20    DIM X(7,2) , Y(7,1) , A(2,7) , B(2,2) , C(2,2), V(2,2) ,
      D(2,1), E(2,1), Z(7,1)
30    DIM P (1,1) , T (1,1), H(1,1), R(1,1), S(1,1), G(1,1), Q (1,2)
      F(1,1), N(3,1)
31    DIM 0(1,1), I(1,1) , K(1,1), U(1,1), L(7,1), M(1,7), J(3,2), W(1,1)
35    N = 2 : REM  NO OF REGRESSION PARAMETERS
40    MAT  READ , X, Y
45    MAT  READ J: REM VALUES TO BE USEDFOR  FORCAST
50    MAT  A = TRN (X): REM  TRASPOST OF MATRIX X
60    MAT  B =  A*X:  REM GIVES N SUMS &
      SUMS OF SQUARES & CROSS – PRODUCTS
70    MAT C= IN V(B): REM USED FOR PARAMETERS
      AND VARIANCE COVARIANCE
80    MAT D = A * Y: REM  GIVES SUMS  AND
      CROSS – PRODUCTS WITH  Y
90    MAT E = C*D: REM COLUMN VECTOR OF PARAMETERS
100   G (1,1) = (D(1,1)/B(1,1)) **2 * B(1,1)
130   MAT Q = TRN (E)
140   MAT  P = Q*D
150   MAT P = P-G
160   A1 = (B(1M1)-1)** (-1)
180   N1 = (N-1) **(-1)
190   MAT  W = (N1) * P
220   MAT M = TRN (Y)
230   MAT T = M * Y
240   MATT = T-G
250   MAT U = INV(T)
260   K (1,1) = B(1,1) – N
265   MAT I = INV (K)
270   MAT  H = T-P
280   MAT S = H * I
```

```
290     MAT O = INV (S)
300     MAT F = O * W
330     MAT R = P*U
340     E1 = S (1,1)
350     MAT V= (E1)*C: REM VARIANCE COVARIANCE MATRIX X
360     S1 = SQR (V(1,1)): REM STANDARD
        DEV.OF FIRST PARAMETER
370     S2 = SQR (V(2,2): REMSTANDARD DEV
        OF SECOND PARAMETER
380     T1 = E (1,1)/ S1: REM T VALUE OF FIRST
390     T2= E(2,1)/ S2: REM T VALUE OF SECOND PARAMETER
400     MAT Z = X*E: REM PREADICTED VALUES
410     MAT  L = Z-Y :  REM RESIDUAL
420     MAT N = J* E: REM FORECAST
490     PRINT, E(1,1);  ,` = تقدير المعلم  أ
500     PRINT, E(2,1); = تقدير المعلم  ب
520     PRINT, R (1,1); ` = معامل التحديد
525     PRINT
527     PRINT
530     PRINT USING 810
540     PRINT USING 820
550     PRINT USING 830
560     PRINT USING 840
570     PRINT USING 850
580     PRINT USING 860, F(1,1), W(1,1), N-1, P(1,1)
585     PRINT
590     PRINT USING 870, S(1,1), K(1,1), H(1,1)
600     PRINT USING 880
610     PRINT USING 890, B(1,1) = 1 ,  T(1,1)
615     PRINT
618     PRINT
620     PRINT,
630     MAT PRINT V
635     PRINT
637     PRINT
```

```
640     PRINT, S1; `   (أ)        الانحراف المعياري للمعلم
650     PRINT
660     PRINT, S2; `   (ب)        الانحراف المعياري للمعلم
665     PRINT
666     PRINT
670     PRINT, T1; `   (أ)               قيمة ت للمعلم
680     PRINT
690     PRINT, T2; `   (ب)               قيمة ت للمعلم
700     PRINT
710     PRINT
715     PRINT USING 900
716     PRINT USING 920
717     PRINT USING 930
718     PRINT USING 950
720     FOR I = 1 TO 7
730     PRINT USING 960, X(I,2), Y(I,1), Z(I,1), L(I,1)
735     PRINT
740     NEXT  I
741     PRINT USING 970
742     PRINT USING 980
743     PRINT USING 990
744     PRINT USING 1000
750     FOR I = 1 TO 3
760     PRINT USING 1010, J (I,2), N (I,1)
770     NEXT  I
810     :                   جدول تحليل التباين
820     : _______________________________________________
830     :  5,1  ف        متوسط      درجات      مجموع      مصدر
840     :                المربعات      الحرية      المربعات   التباين
450     : _______________________________________________
460     :  ****.****   ****.***     **     *****.****  الانحدار (ب) الخطأ
870     :                ****.***        **        *****.****
880     : _______________________________________________
```

```
890          المجموع الكلي *****.*******    **
900          جدول التقدير باستخدام المعادلة
920    : ____________________________________________
930    :     X        Y         التقدير   خطأ التقدير
950    : ____________________________________________
960    :  ****.**   ****.**   ****.****    ***.****
970    :          جدول التنبؤ باستخدام المعادلة
980    :            X              التنبؤ
1000   : ____________________________________________
1010   :           ****.**                    ****.**
1020   DATA 1,10,1,15,1,14,1 1 16, 1,11,1,20,
1030   DATA  1,19,25,35,30,40,24,50,48,
1040   DATA  1,17,1,18,1,25
9999   END
```

المخرجات

تقدير المعلم أ = -5.25

تقدير المعلم ب = 2.749985

معامل التحديد = .9653304

جدول تحليل التباين

مصدر التباين	مجموع المربعات	درجات الحرية	متوسط المربعات	ف 5,1
الانحدار (ب)	635.1875	1	635.187	139.2192
الخطأ الكلي	22.8125	5	4.562	
المجموع الكلي	**658.0000**	**6**		

مصفوفة التشتت

12.87273	-.8147299
-.4147299	5.431531 E – 02

الانحراف المعياري للمعلم (أ) 3.587858

الانحراف المعياري للمعلم (ب) .2330564

قيمة ت للمعلم (أ) -1.463268

قيمة ت للمعلم (ب) 11.79965

جدول التقدير باستخدام المعادلة

X	Y	التقدير	خطأ التقدير
10.00	35.00	22.2498	-2.7502
15.00	35.00	35.9998	0.9998
14.00	30.00	33.2498	3.2498
16.00	40.00	38.7498	-1.2502
11.00	24.00	24.9998	0.9998
20.00	50.00	49.7497	-0.2503
19.00	48.00	46.9997	-1.0003

جدول التنبؤ باستخدام المعادلة

X	التنبؤ
17.00	41.50
18.00	44.25
25.00	63.50

الشروط الواجب توفرها في معادلة الانحدار الخطي:

أهم هذه الشروط، يمكن إيجازها فيما يلي:

أ- اختيار النموذج المناسب، ويتضمن ذلك:

1. أن تكون العلاقة بين المتغيرايت خطية.
2. أن لا يكون هناك متغير ذو علاقة قد تم استبعاده، أو متغير ليست له علاقة أضيف للنموذج.

ب- عدم وجود أخطاء في القياس أثناء جمع البيانات.

ج- عدم وجود ارتباط ذاتي بين المتغيرات Auta Correlation.

د- يجب أن تكون الأخطاء موزعة توزيعاً طبيعياً، كما يجب أن لا يكون بينها ارتباط، وإلاّ تتأثر طردياً أو عكسياً بقيم المتغيرات Homoskedasticity، ويجب أن يساوي وسطهما الحسابي = صفراً.

وللتأكد من توفر هذا الشرط، يتم تنفيذ رسم بياني للأخطاء على القيم التقديرية من النموذج.

برنامج لإجراء العمليات الحسابية الأولية للمصفوفات وبدون استخدام تعليمات المصفوفات

```
10    REM
20    REM  MAT
25    DIM X (7,3), Y(7,1), T(3,7), T(3,7) , E(3,3), H(3,1)
30    REM DIM  X(R,Z), Y(Z,Q), T(Z,R) E(Z,Z), H(Z,Q)
35    REM  X IS THE MATRIX OF INDEP
40    REM Y IS THE MATRIX  OF DEP
60    REM T IS THE TRANSPOSE  OF  X
70    REM E IS THE MATRIX T * X
80    REM H IS THE MATRIX T*X
90    REM TRANSPOSITION PROG.NEXT
100   READ R, Z, Q
110   FOR I = 1 TO R
120   FOR J = 1 TO Z
130   READ X (I,J)
140   NEXT  J
150   NEXT I
160   FOR I = 1 TO R
165   FOR J = 1 TO Q
170   READY (I,J)
180   NEXT J
185   NEXT I
190   FOR I = 1 TO Z
200   FOR J = 1 TO R
210   T (I,J) = X(J,I)
220   NEXT J
230   NEXT I
235   FOR K = 1 TO Z
240   FOR J = 1 TO Z
250   E (K,J)     = 0
260   FOR I = 1 TO R
270   E (K,J) = E (K,J) + T(K,I) * X(I,J)
280   NEXT  I
290   NEXT J
300   NEXT K
310   FOR K = 1 TO Z
325   FOR J = 1 TO Q
330   H (K,J)      = 0
340   FOR I = 1 TO R
350   H (K,J) = H (K,J)+ T(K,I) * Y (I,J)
```

```
360     NEXT I
370     NEXT J
380     NEXT K
390     PRINT MATRIX  (X)
400     FOR I = 1 TO R
410     FOR J = 1 TO Z
420     PRINT X (I,J),
440     PRINT
450     NEXT I
455     PRINT
460     PRINT, MATRIX (T)
470     FOR I = 1 TO Z
480     FOR J = 1 TO R
490     PRINT T (I,J),
500     NEXT J
510     PRINT
520     NEXT
525     PRINT
530     PRINT ` MATRIX (E)'
540     FOR I = 1 TO Z
550     FOR J = 1 TO Z
560     PRINT E (I,J),
570     NEXT J
580     PRINT
590     NEXT I
595     PRINT
600     PRINT, MATRIX (Y)'
610     FOR I = 1 TO R\
620     FOR J = 1 TO Q
630     RPINT Y (I,J),
640     NEXT  J
650     PRINT
660     NEXT I
665     PRINT
670     PRINT, MATRIX (H)'
680     FOR I = 1 TO Z
690     FOR J = 1 TO Q
700     PRINT H (I,J),
710     NEXT J
720     PRINT
730     NEXT I
```

```
740       PRINT
750       DATA 7M3M1
760       DATA 1,23,9,1,28,12,1,21,12,1,23,22,1,30
               16,1,32,18,1,25,23,3,6,5,8
770       DATA  10, 15,9
999       END
```

المخرجات

MQTRIX (X)

1	23	9
1	28	12
1	21	12
1	23	22
1	30	16
1	32	18
1	25	23

MATRIX (T)

1	1	1	1	1
1	28	21	23	30
23	12	12	22	16
9				
23				

MATRIX (E)

7	182	112
182	4832	2932
112	2932	1962

MATRIX (Y)

3
6
5
8
10
15
9

MATRIX (H)

56
1531
972

الوحدة العاشرة

الارتباط غير الخطي
(*Non-Linear Correlation*)

أو توفيق المنحنيات Curve Fitting

في كثير من الأحيان تكون طبيعة العلاقة بني المتغيرين غير خطية.

في بعض الأحيان يمكننا التعبير عن العلاقة غير الخطية بين المتغيرين بعلاقـة خطيـة، فمـثلا بـدلاً من توفيق منحنى لوصف العلاقة

$$Y = \frac{1}{A + BX}$$

يمكننا استخدام مقلوب المتغير التابع

وليكن $Y^* = \frac{1}{Y}$ وتوقيف خط مستقيم للعلاقة

$$Y^* = \frac{A + BX}{Y}$$

$$Y^* = A + BX$$

وذلك بطريقة المربعات الصغرى باستخدام المعادلات:

$$\sum Y^* = NA + B\sum X$$

$$\sum XY^* = A\sum X + B\sum X^2$$

وهكذا يمكننا تقدير قيمة الثوابت A,B ومنها نقدر العلاقة المطلوبة:

$$\hat{Y} = \frac{1}{A + BX}$$

بيـنما إذا كانـت العلاقـة بـين الظـاهرتين يمكـن وصـفها بـالمنحنى Y = ABX فإنـه باسـتخدام اللوغاريتمات، نستطيع تحويلها إلى علاقة خطية:

LOG Y = Log A + Log B (X)

$$\hat{Y}* = A* + B*X$$

حيث تدل العلاقة (*) على استخدام القيمة المعدلة بعد أخذ اللوغاريتم.

مثال:

إذا علمت أن Y = ABX، فاحسب معادلة انحدار $\frac{Y}{X}$ من البيانات التالية:

X:	1	2	3	4	5	6	7
Y:`	304	344	393	457	548	670	882

الحل:

يمكن تحويل معادلة المنحنى إلى معادلة خطية باستخدام اللوغاريتمات فتصبح:

LogY = A* + B+X

ونحسب القيم المطلوبة لحل المعادلات الآتية أو للتعويض في القـوانين مبـاشرة كـما يتضـح مـن الجدول التالي:

X	Y	Y^2	XY^2	X^2
1	304	2.483	2.483	1
2	341	2.533	5.066	4
3	393	2.594	7.782	9
4	457	2.660	10.640	16
5	548	2.739	13.695	25
6	670	2.826	16.956	36
7	882	2.945	20.516	49
28		**18.780**	**77.237**	**140**

وبحساب المتوسط الحسابي للمتغيرين نجد أنهم:

$$\overline{X}1 = 4, \qquad Y^* = 2.683$$

وبالتعويض في القانون:

$$B = \frac{N\sum XY - (\sum X)(\sum Y)}{N\sum X^2 - (\sum X)^2}$$

$$= \frac{7(77.237) - 28(19.780)}{7(140) - (28)^2} = 0.076$$

$$A = \overline{Y}' - B'\overline{X} = 2.683 - 0.076\ (4) = 2.379$$

منها نحسب المعالم الأصلية، ونجد أن:

A = 239

B = 1.19

أي أن المعادلة المطلوبة هي:

$$\hat{Y} = (239)\ (1.19)X$$

معادلة النهاية العظمى واحدة إلى الأعلى أو لأسفل فيكون له نهاية صغرى واحدة.

وإن هذه المعادلة تكون على الصورة التالية $\hat{Y} = A + BX + CX^2$

ويمكننا استخدام طريقة المربعات الصغرى لتقدير المعالم لمعادلة الانحدار غير الخطية أيضاً. يتم ذلك بحل المعادلات الطبيعية التالية:

$$\sum Y = NA + B\sum X + C\sum X^2$$

$$\sum XY = A\sum X + B\sum X^2 + C\sum X =$$

$$\sum X2Y = A\sum X^2 + B\sum X^3 + C\sum X^4$$

وهكذا نحصل على معادلة انحدار $\frac{Y}{X}$

مثال:

احسب معادلة انحدار Yعلى X، علماً بأن العلاقة بينهما من الدرجة الثانية:

X:	4	2	5	2	3
Y:	2	1	3	2	1

الحل:

المعادلة المطلوبة تأخذ الصورة $\hat{Y} = A + BX + CX^2$.

وللحصول على قيم المعالم أي الثوابت، نحتـاج إلى حـل المعـادلات الـثلاث السـابقة، ونحتـاج إلى حساب القيم الواردة في الجدول التالي للتعويض فيهم.

X	Y	Y^2	X^3	X^4	X^2Y	Y^2
4	2	16	64	256	32	4
2	1	4	8	16	4	1
5	3	25	125	625	75	9
2	2	4	8	16	8	4
3	1	9	27	81	9	1
16	**9**	**58**	**232**	**994**	**128**	**19**

$9 = 5A + 16B + 58C$

$32 = 16A + 58B + 232C$

$128= 58A + 232V + 9948C$

وبحل هذه المعادلات، نحصل على:

$A = 4.2$ $B = 2.09$ $C= 0.37$

وبالتالي المعادلة المطلوبة هي:

$\hat{Y} = 4.2 – 2.09X + 0.37X^2$

1/10 قياس الارتبـــاط

في حالة الارتباط غير الخطي، يمكن استخدام التشتت حول خط الانحـدار للحصـول عـلى معيـار للارتباط.

يقاس التشتت حـول منحنـى الانحـدار، باسـتعمال مربـع انحرافـات القـيم الفعليـة عـن القـيم المحسوبة باستخدام معادلة الانحدار، وتطبق الصورة التالية لحساب الخطأ المعياري.

$$S\hat{Y} = \frac{\sqrt{\sum (Y - \hat{Y})^2}}{N}$$

وتسهيلاً للعمليات الحسابية، واختصاراً للخطوات، نستخدم إحدى الصور التالية:

$$S\hat{Y} = \frac{\sum Y^2 - A\sum Y - B\sum XY - C\sum X^2 Y}{N}$$

$$S\hat{Y} = SY\sqrt{1 - (I)^2}$$

حيث:

(I) تسمى دليل الارتباط Correlation index، وهي مقيـاس للارتبـاط غـير الخطـي، مهـما كانـت درجة العلاقة بين الظواهر محل الدراسة. ومن المعادلة، نرى أن دليل الارتباط يمكنحسابه بالقانون:

$$I = \sqrt{1 - \frac{\sigma\hat{Y}}{\sigma Y}}$$

حيث $\sigma\hat{Y}$ و σY هي تباين القيم المقدرة، والقيم الفعلية على التوالي.

- يتساوى كل من r,I معامل الارتباط إذا كانت العلاقة بين المتغيرين Y,X خطية، حيث يكون تعريف الخطأ المعياري واحد في الحالتين. ويختفي العنصر الأخير من قانونه الثاني أي $\left(C\sum X^{2}Y\right)$.

- تزيد قيمة (I) عن (r) إذا كانت العلاقة غير خطية، وكان توفيق منحنى بين المشاهدات سيعطي نتائج أفضل، لأن التشتت حول المنحنى سيكون أقل من التشتت حول خط مستقيم، وبالتالي سيكون الخطأ المعياري حول المنحنى أقل، وستزيد قيمة (I) عن (r). بعبارة أخرى "إن دليل الارتباط لما يساوي أو يزيد عن معامل الارتباط".

- فدليل الارتباط يزيد عن معامل الارتباط في حالة الانحدار غير الخطي على اختلاف أنواعه. حيث أن قانون الخطأ المعياري لا يختصر إلى القانون المستخدم في حالة الارتباط الخطي. لذلك يعتبر دليل الارتباط أعم وأدق من معامل الارتباط في قياس العلاقة بني متغيرين.

مثال:

احسب دليل ا لارتباط من الأرقام المستخدمة بالمثال السابق مباشرة:

$$\sigma\hat{Y} = \frac{\sum Y^{2} - A\sum Y - B\sum XY - C\sum X^{2}Y}{N}$$

$$= \frac{19 - \left[(4.02)\,(0.9)\right] - \left[(-2.09)\,(32)\right] - \left[(37)(129)\right]}{5}$$

$$= 0.14$$

$$\sigma Y = \frac{\sum Y^{2}}{N} - \left(\frac{\sum Y}{N}\right)^{2}$$

$$= \frac{19}{5} - \left(\frac{(9)}{5}\right)^{2}$$

$$= 0.56$$

$$I = \sqrt{1 - \frac{0.14}{0.56}} \quad = \quad 0.866$$

- هذا المثال يحدد لنا درجة العلاقة بين المتغيرين.
- في حين أن الطريقة الأفضل هي حساب معامل الارتباط ودليـل الارتبـاط ودليـل الارتبـاط إذا لم طبيعة العلاقة معروفة.
- إذا وجدنا أن دليل الارتباط أكبر دل ذلك على العلاقة بين المتغيرين غير خطية.
- يلاحظ أن دليل الارتباط لا يقل أبداً عن معامـل الارتبـاطن وبالتـالي لا يمكـن القـول أن العكـس صحيح.
- في حالة تساوي المقياسين، فإننا نلجأ إلى اختيارات أخرى لتحديد العلاقة بني المتغيرين.

2/10 الارتباط المتعدد والارتباط الجزئــــي

Multiple & Partial Correlation

إذا رغبنا دراسة العلاقة بين الطلب على سلعة والظـواهر الأخـرى مجتمعـة، نسـتخدم الارتبـاط المتعدد.

عند قياس درجة العلاقة بين ثلاث ظواهر أو أكثر، تسمّى العلاقة بارتباط متعدد، فالمتغير التـابع معتمد على عدد من المتغيرات المستقلة.

دراسة العلاقة بين التغير في عدد من المتغيرات المسـتقلة والتغـير في ظـاهرة معينـة، أي أن: Y= F(X,W,Z,…).

قد تكون Y مثلاً كمية الإنتاج الزراعي، ويعتمد التغير في كمية السـماد المسـتخدم وكميـة الميـاه المستخدمة في الري، قد تأخذ الدالة صورة خطية فتكون: $\hat{Y} = A + BX + SW$.

وقد تكون طبيعة العلاقة بين الظواهر محل الدراسة غير خطية، بصورة

$$\hat{Y} = BX + SW^2 + CZ^3$$

لتقدير قيم الثوابت في المعادلة:

$$\hat{Y} = A + BX + CW$$

نستخدم المعادلات:

$$\sum Y = NA + B\sum X + C\sum W$$

$$\sum XY = A\sum X + B\sum X^2 + C\sum XW$$

$$\sum XY = A\sum W + B\sum WX + X\sum W^2$$

وبحل تلك المعادلات، نحصل على قيم المعالم A,B,C لحساب الخطأ المعياري، يؤخذ الجذر التربيعي لمربع انحرافات القيم الفعلية عن القيم المتنبأ بها، باستخدام القانون:

$$S\hat{Y} = \sqrt{\frac{\sum (Y - \hat{Y})^2}{N}}$$

حساب تباين التقديرات، يتم بالقانون:

$$\sigma Y = \frac{\sum Y^2 - A\sum Y - B\sum XY - C\sum WY}{N}$$

ثم يكون معامل الارتباط المتعدد بين Y، وكل من W,X

$$ry, xw = \sqrt{1 - \frac{\sigma\hat{Y}}{\sigma Y}}$$

حيث σY تباين القيم الفعلية للمتغير

$\sigma\hat{Y}$ تباين القيم المقدرة.

ويقيس معامل الارتباط المتعدد العلاقة بين التغير في المتغير التابع y والتغير في المتغيرات المستقلة (W,X) مجتمعة.

الارتباط الجزئي **Partial Correlation**

الارتباط الجزئي، هو الارتباط الذي يقيس العلاقة بـين المتغـير التـابع وأحـد المتغـيرات المسـتقلة، حين تبقى المتغيرات المستقلة ثابتة عند مستوى معيّن.

إذا رغب الباحث في دراسة العلاقة بين Y (المتغير التابع) والمتغير المستقل (x) فقط بفـرض بقـاء المتغير المستقل (w) ثابت، فتتحول معادلة ا لارتباط المتعدد

$Y = A + BX + CW$

إلى الصورة التالية $Y' = A' + B'X + D'$ حيث D' كمية ثابتة.

وبالمثل إذا أخذنا معادلة انحدار $\frac{X}{Y}$ مع بقاء W على حالها نحصل على

$X' = A' + B'Y + D'$ حيث D' كمية ثابتة.

وبأخـذ الجـذر التربيعـي لحاصـل ضرب معـاملي الانحـدار، نحصـل عـلى معامـل الارتبـاط $RYX.W = \sqrt{B.B'}$

الوحدة الحادية عشر

السلاسل الزمنية
Time Series

السلسلة الزمنية هي: مجموعة من القيم لظاهرة ما في فترات زمنية متعاقبة، طبقاً لأزمنة حدوثها.

فالسلسلة تحتوي على متغيرين، إحداهما تابع (Y) مثلاً والآخر هو (T) الزمن كمتغير مستقل. وتكون العلاقة محل الدراسة هي $Y=F(T)$.

هناك سلسلة زمنية للأسعار، للإنتاج، الصادرات، للسكان.

من الضروري مراعاة استخدام نفس الوحدات القياسية، ونفس طريقة القياس للظاهرة، وإلاّ انعدمت إمكانية المقارنة بين قيم الظاهرة في فترات مختلفة.

يمكن تمثيل السلسلة الزمنية، بنقطة تتحرك عبر الزمن، تاركة خط منكسر يمثل قيم الظاهرة محل الدراسة.

1/11 عناصر السلسلة الزمنية

يعني تحليل السلاسل الزمنية، دراسة تلك المؤثرات للتعرف على ما تعرضت له الظاهرة في الماضي، واستخدام هذه المعلومات للتنبؤ بقيمة الظاهرة محل الدراسة بالتغير في سلسلة أخرى أو أكثر.

هناك أربعة مؤثرات تطرأ على الظواهر عبر الزمن هي:

1. المؤثرات الخارجية (الاتجاه العام).
2. المؤثرات الموسمية.
3. التغيرات الدورية.
4. التغيرات العرضية.

1/1/11 الاتجاه العام (المؤثرات الخارجية)

هذه المؤثرات تجعل الظاهرة تتبع مجرى ثابت عبر فترة طويلة من الزمن (للزمن)، ويكون هـذا الاتجاه تصاعدي إذا ما اتجهت قيم الظـاهرة إلى التزايـد، أمـا إذا مالـت قـيم الظـاهرة للتنـاقص، فـإن الاتجاه العام تناقصي.

وقد يكون الاتجاه العام خطي أي يمكن تمثيله بمستقيم أو غير خطـي، أي يمكـن تمثيلـه بمنحنـى من الدرجة الثانية أو أعلى.

2/1/11 التغيرات الموسمية

هي تقلبات روتينية تطرأ على الظاهرة خلال الفترة الزمنية محل الدراسة، تتكـرر بانتظـام، فقـد تكون يومية، أسبوعية، شهرية، ربع سنوية، درجة الحرارة في أي يوم تبدأ منخفضـة ثـم ترتفـع وترجـع إلى الانخفاض في المساء، حجم المبيعات يرتفع في بداية كل شهر.

3/1/11 التغيرات الدورية

التغيرات الدورية، هي تقلبات على مدى أطول من التغيرات الموسمية التي تتكرر خـلال السـنة أكثر من مرة واحدة. إلا أن التغيرات الدورية تحركها لفترة أقل طولاً من فترة الاتجاه العام.

التغيرات الدورية تقلبات طويلة الأجل حول خط الاتجاه العام.

والتغيرات الدورية تعكس تتابع فترات الكساد والرواج التي يمر بها الاقتصاد القومي.

4/1/11 التغيرات العرضية

هي تقلبات تعود على ظروف طارئة فجائية لا يمكن التنبؤ بها مقدماً، فهي لا تحـدث طبقـاً لقاعـدة معينة، بل عوامل فجائية كالحروب والفيضانات وغيرها من الكوارث الطبيعية التـي تـأثر في الظـاهرة محـل الدراسة، وقد تتكون بعد ذلك أو لا تتكرر.

1/1/11 تحديد الاتجاه العام

هناك عدة طرق تستخدم لتعيين الاتجاه العام لسلسلة زمنية تتفاوت فيما بينها من حيث دقتها في تحديد خط الاتجاه العام، ومن أهمها:

1. التمهيد اليدوي.
2. طريقة الوسط النصفي.
3. طريقة المتوسطات المتحركة.
4. طريقة المربعات الصغرى.

1/1/1/11 التمهيد باليد "الرسم البياني"

تعرض السلسلة الزمنية بياناً، بأخذ المحور الأفقي لتمثيل الزمن، والمحور الرأسي لتمثيل الظاهرة محل الدراسة، ثم نحدد جميع النقاط بإحداثياتها الأفقي والرأسي، ونوصلها، فنحصل على المنحنى التاريخي للظاهرة خلال الفترة المدروسة، ويجب أن نختار مقياس الرسم المناسب. ثم يحاول الباحث أن يمهد الخط أو المنحنى الذي يمر بين أكبر عدد من المشاهدات لتحديد ما يبدو أنه التقلبات الطويلة الأجل أي الاتجاه العام.

إذا كان الاتجاه العام خطياً فإنه يمكننا تحديد المعادلة التي توصف الاتجاه العام بحساب ميل الخط الممهد وتحديد الجزء المقطوع من المحور الرأسي.

إن هذه الطريقة تمتاز بالبساطة، إلا أنه لا يمكن الاعتماد على دقة النتائج، حيث أنها تتوقف على خبرة ومران الباحث، وللعامل الشخصي تأثير كبير من النتائج، لذا يفضل الاعتماد على الطرق الأخرى.

2/1/1/11 طريقة الوسط النصفي

في هذه الظاهرة نقسم السلسلة إلى قسمين متساويين (إذا كان عدد السنوات فردي، نهمل قيمة الظاهرة، في السنة الوسطى) ثم نحسب المتوسط الحسابي لكل قسم

وتحدد مكانهما على الرسم. كـل يقابـل السـنة الوسـطى بالنسـبة لكـل قسـم، ثـم نمهـد الخـط المستقيم الذي يمر بالوسطين، فيكون خط الاتجاه العام.

مثال:

أوجد خط الاتجاه العام للسلسلة الزمنية الآتية باستخدام طريقة الوسط النصفي:

T:1980	1971	1972	1973	1974	1975	1975	19777	1978	1979
Y: 3	5	7	10	15	14	15	17	19	20

الحل:

حيث أنه لدينا عشر مشاهدات فنقسم السلسلة إلى مجموعتين تحتـوي كـل منهـا عـلى خمـس سنوات ثم نحسب المتوسط الحسابي لكل منها:

$$=\frac{3+5+7+10+15}{5}=8$$

$$=\frac{14+15+17+19+20}{5}=17$$

يقابل الوسط الحسابي لكل مجموعة نقط الوسط في تلك المجموعة. وحيـث أن عـدد السـنوات فردي، فسوف يقابله السنة الوسطى، أي أن الوسط الحسابي (8) يقابل 1972.

والوسط الحسابي (17) للمجموعة الثانية سوف يقابل 1977 بتحديد تلـك النقطتـين عـلى الرسـم البياني، وبتوصيلهم نحصل على خط الاتجاه العام (أما إذا كان عدد السنوات زوجي في كل قسـم، فـإن الوسط الحسابي سيقع بين القيم الأصلية، ولن يقابل سـنة بـذاتها، بـل سـيقع بـين سـنتين مـن سـنوات السلسلة أي سيقابل يناير من السنة الأخيرة، حيث أن سنوات السلسلة تبدأ من يوليو).

يمكننا حساب معادلة خط الاتجاه من المعلومات السابقة أيضاً، حيث أن قيمة الظاهرة قد ارتفعت بمقدار 9 وحدات (8-17) خلال فترة الخمس سنوات 1972-1977، أي أن مقدار التغير السنوي 1.8 وهو ميل خط الانحدار، أي أن:

ميل خط الاتجاه العام = $\frac{\text{الفرق بين الوسطين الحسابيين}}{\text{الفرق الزمني}}$

أما الجزء المقطوع من المحور الرأسي فيعتمد على نقطة الأصل، حيث نستخدم الوسط الحسابي المقابل لواحدة من نقطتي الوسط، فإذا أخذت سنة 1972 كنقطة الأصل، فإن معادلة الاتجاه العام هي:

(باستخدام سنة 1972 كنقطة أصل) $\hat{Y} = 8 - 1.8T$

"ومن هذه المعادلة يمكن حساب القيم الاتجاهية للظاهرة، أي تلك التي تأخذها الظاهرة لو وقعت على هذا الخط".

إذا رغبنا في حساب القيمة الاتجاهية للظاهرة سنة 1975، فنعوض عن الزمن بالفرق سواء أكان موجباً أو سالباً بين السنة المطلوبة وسنة الأساس

$$\hat{Y}_{(1975)} = 8 - (1.8)(3) = 13.4$$

أما إذا استخدمنا 1977 كسنة أساس (أي نقطة الأصل)، فإن المعادلة تكون (سنة الأساس هي 1977)

$$\hat{Y} = 17 - (1.8)T$$

من الضروري تحديد نقطة الأصل للمعادلة المعطاة، حتى يمكن حساب الفرق بين نقطة الأصل والسنة المطلوبة، فالقيمة الاتجاهية لسنة 1975 باستخدام المعادلة الثانية بالتعويض عن قيمة T (2-).

$$\hat{Y} = 17 - (1.8)(-2)$$

ويعاب على طريقة الوسط النصفي، بأنها تتأثر بالقيمة المتطرفة، سواءً الشديدة الارتفاع أو الشديدة الانخفاض، مما يؤثر على قيمة الوسطين الحسابين المحسوبين، وبالتالي يؤثر في ميل خط الانحدار، وفي دقة المعادلة المحسوبة، خاصة إذا كانت البيانات أو السنوات المتوفرة محدودة.

وللتغلب على هذا العيب، يمكننا حساب الوسيط بدلاً من الوسط الحسابي لكل مجموعة (تسمى هذه الطريةق بالوسيط النصفي)، ثم الاستمرار بالخطوات كالعادة.

إن طريقة الوسط النصفي تستخدم لتمهيد خط اتجاه عام. ولا يمكن استخدامها لتوفيق منحنى، إذا ما كان هذا أفضل من الخط المستقيم في وصف الاتجاه العام للظاهرة محل الدراسة.

3/1/1/11 طريقة المتوسطات المتحركة

تعتبر طريقة المتوسطات المتحركة امتداد للمتوسط النصفي بدلاً من أخذ متوسطين حسابين، الأفضل أخذ عدد أكبر من المتوسطات، حتى نحقق مستوى أعلى من الدقة في تحديد خط الاتجاه العام.

يمكن تقسيم السلسلة الزمنية إلى عدة مجموعات مستقلة، نحسب المتوسط الحسابي لكل منها، ولكن الأفضل أن نستعمل مجموعات متداخلة، نحسب لكل منها المتوسط الحسابي، ثم نسقط القراءة الأولى من المجموعة، ونضيف قيمة جديدة، بحيث يكون لدينا نفس العدد من القراءات في كل مجموعة، ونحسب الوسط الحسابي للمجموعة الثانية، ثم نعيد الكرة مرة أخرى بأن نسقط القراءة الأولى من المجموعة الثانية ونضيف قيمة جديدة ونحسب المتوسط الحسابي... وهكذا.

مثال:

احسب المتوسطات المتحركة على أساس ثلاث سنوات من المثال السابق.

الحل:

المتوسط المتحرك الأول = $\frac{3+5+7}{3} = 5$

المتوسط المتحرك الثاني = $\frac{5+7+10}{3} = 7.33$

كما هو وارد في الجدول التالي:

T	Y	المجموع المتحرك لثلاث سنوات	المتوسط المتحرك	
1970	3	-	-	لا يوجد متوسط مناظر للقيمة الأولى
1971	5	15	5	
1972	7	22	7.33	
1973	10	32	10.66	
1974	15	39	13.00	
1975	14	44	14.66	
1976	15	46	15.33	
1977	17	51	17.00	
1978	14	56	18.66	
1979	20	-	-	لا يوجد متوسط مناظر للقيمة الأخيرة

- ويلاحظ أنه لو كان عدد الفترات التي يحسب على أساها المتوسط الحسابي (فردياً)، فإن المتوسط المحسوب يمثل القيمة الاتجاهية للظاهرة محل الدراسة في النقطة الوسطى لكل مجموعة.
- أما إذا كان عدد الفترات (زوجي)، فإننا نلجأ إى ما يمّى "بمركزه المتوسط المتحرك".
- لا يوجد متوسط مناظل للقيمة الأولى والقيمة الأخيرة، وعموماً إذا كان طول فترة التجميع يساوي (N) فإننا نفقد (N-1) متوسط متحرك دائماً.
- أما إذا حسبنا المتوسطات المتحركة لفترة تجميع مكونة من عدد زوجي من السنوات، فإن المتوسطات المتحركة ستقع بين القيم الأصلية، ولن تقابل إحداها،

لذا فإننا نقوم بحساب متوسط متحرك على أساس فترتين للمتوسـطات المتحركـة الأولى، وسـوف تقابل المتوسطات الجديدة الفترة الأخيرة، وتسمى هذه العملية بمركزة المتوسطات.

مثال:

احسب المتوسطات المتحركة على أساس أربع سنوات مستخدماً بيانات المثال السابق.

الحل:

الطريقة الأولى سوف نصنف الخطوات المطلوبة في الجدول التالي:

T	Y	المجموع المتحرك 4 سنوات	المتوسط المتحرك	مجموع متحرك لمتوسطين	القيمة المركزة
1970	3				
1971	5	25	6.25	15.50	7.75
1972	7	37	9.25	20.75	10.37
1973	10	46	11.50	25.00	12.50
1974	15	54	13.50	28.75	14.37
1975	14	61	15.25	30.25	15.12
1976	15	60	15.00	33.50	16.75
1977	17	66	18.50		
1978	14				
1979	20				

في هذه الطريقة حسبنا مجموع أربعة قراءات، ثم أهملنا الأولى، وأضفنا قراءة جديدة للحصول على المجموع المتحرك... وهكذا.

إن عدد المتوسطات المتحركة $3+5+7+10 = 25$

$5+7+10+15 = 37$

يقل عن القراءات بمقدار 3 وهو (N-1).

1/3/1/1/11 المتوسط المتحرك المرجح

قد يحصل الباحـث مقـدماً عـلى أوزان حسـابية، يمكـن اسـتخدامها في عكـس الأهميـة النسـبية للقراءات، وبالتالي حساب متوسط مرجح.

إذا فرضنا أن الباحث يرغب في حساب متوسط متحرك على أساس ثلاث فترات، وكانت الأوزان الحسابية هي 1,2,2 فإننا نحسب المتوسط المتحرك المرجح بضرب المشاهدة في الوزن المعطى لها، ثم نقسم على مجموع الأوزان، ثم نلغي القراءة الأولى، ونضيف أخرى ونضرب في نفس الأوزان ونقسم على مجموعها، وهكذا.

أي أن كل قراءة تضرب في كل الأوزان طبقاً لمكانها عند حساب المجموع المتحرك.

مثـال:

باستخدام نفس السلسلة الزمنية السابق والأوزان 1,2,2 احسب المتوسط المتحرك المرجح.

الحل:

نلاحظ أن المجموع المتحرك الأول يحسب كالتالي:

(3) (1) + (5)(2) + (7)(2)= 27

أما الثاني، فسيكون (5) (1)+(7)(2) + (10)(2) = 39

وهكذا بالنسبة للقية التي تعرض في جدول كالتالي:

المتوسط المتحرك المرجح	المجموع المتحرك المرجح	Y	T
-	-	3	1970
5.4	27	5	1971
7.8	39	7	1972
11.4	57	10	1973
13.6	68	15	1974
14.6	73	14	1975
15.6	787	15	1976
17.4	87	17	1977
19.0	95	14	1978
-	-	20	1979

- نلاحظ أن طريقة المتوسطات المتحركة سواءً المرجحة أو غير المرجحة لا تعني معادلة خط الاتجاه العام، بل القيم الاتجاهية فقط.

- كما أنها تحتاج معرفة أنسب عدد للفترات المطلوبة لحساب المتوسطات المتحركة.
- عدد القراءات يتزايد مع طول فترة التجميع، وهذا عيب هام، خاصة إذا كانت عدد القراءات المتاحة صغيراً.

4/1/1/11 طريقة المربعات الصغرى

إن طريقة المربعات الصغرى، هي أفضل الطرق وأكثرها انتشاراً، لتحديد خط الاتجاه العام لسلسلة زمنية على اساس أن مجموع مربعات انحرافات القيم المشاهدة عن الخط الممهد يكون أصغر ما يمكن، ومجموع الانحرافات عن الخط يكون مساوياً للصفر.

مثال:

احسب معادلة خط الاتجاه العام للسلسلة الزمنية التالية باستخدام طريقة المربعات الصغرى.

T	Y	T	Y
1951	19	1956	25
1952	21	1957	25
1953	28	1958	27
1954	25	1959	31
1955	24		

وسوف نعطي السنة الأولى بالسلسلة القيمة صفر، ونعتبرها سنة الأساس.

T	X	Y	XY	X^2
1951	0	19	0	0
1952	1	21	21	1
1953	2	23	46	4
1954	3	25	75	9
1955	4	24	96	16
1956	5	25	125	25
1957	6	25	150	36
1958	7	27	189	49
1959	8	31	248	64
	36	**220**	**950**	**204**

$$B = \frac{N\sum XY - \left(\sum X\right)\left(\sum Y\right)}{N\sum X^2 - \left(\sum X\right)^2}$$

$$A = \overline{Y} - B\overline{X}$$

$$B = \frac{(9) - (950) - (36)\,(220)}{(9) - (204) - (36)\,(36)} = 1.166$$

$$A = 24.444 - (1.166)\,(4) \qquad = 19.78$$

$$\hat{Y} = 19.78 + 1.166\,(X)$$

"علماً بأن سنة الأساس 1951، إن وحدة قياس الزمن هي السنة".

ويمكننا استخدام المعادلة في التنبؤ بقيمة الظاهرة في عام 1961 مثلاً أن نعـوض في المعادلـة عـن (X) بالفرق بين السنة المطلوبة وسنة الأساس أو نقطة الأصل التي أعطيت القيمة صفر بالنسبة للزمن.

$$\hat{Y}_{1961} = 19.78 + (1.166)\,(10) = 31.44$$

وبالطبع القيمة الفعلية في هذه السنة، قد تختلف عن القمية المقـدرة أو المتنبـأ بهـا، والأخـيرى (المتنبأ بها) توضح لنا قيمة الظاهرة لو كانت تحت تـاثير الاتجـاه العـام للظـاهرة فقـط، ولـذا تسـمّى القمي المقدرة بالقيم الاتجاهية.

يمكننا تبسيط العمليات الحسابية، بنقل سنة الأسـاس عـلى منتصـف السلسـلة الزمنيـة، بحيـث يكون مجموع قيم X مساوياً للصفر، وحيث أن عدد السنوات المتوفرة بالسلسلة فردي، فنأخـذ السـنة "الوسطى" كسنة أساس، أي نقطة الصفر، وتكـون السـنوات السـابقة لهـا ...3-,2-,1-، وتكـون الفـترات اللاحقة لسنة الأساس هي الأول والثانية ...,3,2,1...

كما يتضح لنا من حل نفس المثال السابق، بانحرافات عن السنة الوسطى.

T	X	Y	XY	X^2
1951	-4	19	-76	16
1952	-3	21	-63	9
1953	-2	23	-46	4
1954	-1	25	-25	1
1955	0	24	0	0
1956	1	25	25	1
1957	2	25	50	4
1958	3	27	81	9
1959	4	31	124	16
	0	**220**	**70**	**60**

حيث أن مجموع X يساوي الصفر، لذا تختصر قانوني حساب الثوابت إلى:

$$B = \frac{\sum XY}{\sum X^2} \quad , \quad A = \frac{\sum Y}{N}$$

$$B = \frac{70}{60} = 1.166 \quad , \quad A = \frac{220}{9} = 24.444$$

(السنة الأساس سنة 55، وحدة القياس بالسنة)

$$\hat{Y} = 24.444 + 1.66X$$

وللتنبؤ بقيمة الظاهرة في سنة 1961 مثلاً، نعوض عن (X) بالقيمة (6) فقط، حيث أن سنة الأساس قد تغيرت، ولذا تعتبر البيانات المذكورة بين قوسين هامة للغاية.

$$\hat{Y}_{1961} = 24.444 + (1.166)\ (6) = 31.44$$

أما إذا كان عدد السنوات بالسلسلة زوجي، فيمكن باستخدام الطريقة المطولة، واعتبار السنة الأولى بالسلسلة هي سنة الأساس، والتعويض عنها بصفر بالنسبة للمتغير X. أي الزمن، واتباع نفس الخطوات.

أما إذا استخدمنا الطريقة المختصرة، فإن النقطة الوسطى في السلسلة ستقع بين السنتين الوسطين، أي في يناير من السنة الأخيرى، ويكون الفرق بين نقطة الوسط وقيمة الظاهرة في السنة التالية لها هو نصف سنة، أي الفرق بين يناير ويوليو، ثم يزيد هذا الفرق بمقدار واحد في السنة التالية... وهكذا لبقية السنوات.

مثال:

احسب معادلة خط انحدار السلسلة الزمنية التالية:

T	Y
1960	3
1961	5
1962	7
1963	10
1964	12
1965	14
1966	15
1967	17

الحل:

نلاحظ أن نقطة المنتصف في هذه السلسلة هي النقطة الوسطى بين سنة 1963 وسنة 1964، وحيث أن قيمة الظاهرة معطاة لنا في يوليو من كل سنة، فالنقطة الوسطى تكون يناير سنة 1964، ونحسب الانحرافات الزمنية عن هذه النقطة فتكون لنا المتغير X.

X	Y	XY	X^2
-3.5	3	-10.5	12.25
-2.5	5	-12.5	6.25
-1.5	7	-10.5	2.25
-0.5	10	-5.0	.25
0.5	12	6.0	.25
1.5	14	21.0	2.25
2.5	15	36.5	6.25
3.5	17	39.5	12.25
	83	85.5	

$$A = \frac{\sum Y}{N} = 10.37$$

$$B = \frac{\sum XY}{\sum X^2} = 2.03$$

ومعادلة الانحدار هي:

$$\hat{Y} = 10.34 + 2.03X$$

"نقطة الأصل يناير 1964، وحدة قياس الزمن السنة".

إذا رغبنا مثلاً في التنبؤ بقيمة الظاهرة في سنة 1968، فإننا نعني بـذلك يوليـو مـن هـذه السـنة، ولذا يكون الفرق بينها وبين نقطة الأصل (4.5) سنة.

$$\hat{Y}_{1978} = 10.37 + (2.03)\ (4.5) = 19.53$$

ويمكننا التخلص من الكسور العشرية وتبسيط العمليات الحسابية بأن نأخذ وحدة قياس الـزمن على أنها نصف سنة وليست سنة كاملة.

ويتم ذلك بضرب الانحرافات (X) في الرقم (2) كما في الجدول التالي:

بالنسبة للمثال السابق

X	Y	XY	X^2
-7	3	-21	49
-5	5	-25	25
-3	7	-21	9
-1	10	-10	1
1	12	12	1
3	14	42	9
5	15	75	25
7	17	119	49

$$A = \frac{83}{8} = 10.7$$

$$B = \frac{171}{168} = 1.017$$

أي أن معادلة الانحدار هي:

$$\hat{Y} = 10.37 + 1.07X$$

"نقطة الأصل يناير سنة 1964، وحدة قياس الزمن نصف السنة".

فإذا رغبنا في تقدير قيمة الظاهرة في سنة 1968 فإن الفرق الزمني بين يوليو 1968 ويناير 1964 هو (4.5) سنوات لكننا نعوض في المعادلة السابقة بالقيمة (9)، وهي عدد الوحدات القياسية للزمن أي ضعف عدد السنوات حيث أن وحدة القياس هي نصف السنة.

$$\hat{Y}_{1978} = 10.37 + (1.017)\ (9) = 19.53$$

وهي بالطبع نفس القيمة التي حصلنا عليها بالطريقة السابقة.

1/2/1/11 أثر الموسم (التغيرات المسمية)

قد تطرأ على الظاهرة محل الدراسة تغيرات منتظمة من حيث توقيفها حدوثها كل عام. تختلف تلك التغيرات المنتظمة طبقاص لطبيعة الظاهرة محل الدراسة، فقد تتكرر يومياً، مثل ازدحام الحافلات العامة بالركاب بأوقات الدوام الرسمي للذهاب للعمل، والخروج من العمل، أو تتكرر عند الخروج للنزعة أيام الجمعة من فصل الربيع كما قد تكون شهرية مثل زبائن المصرف في الأيام الأولى من كل شهر... وهكذا.

الهدف من دراسة أثر الموسم هو الوقوف على حجم تلك الآثار، أما بتخليص الظاهرة محل الدراسة منها، أو لأخذها في الحسبان حين اتخاذ قرارات تتعلق بالظاهرة.

تقوم الفكرة الأساسية على اعتبار أن السلسلة الزمنية، حاصل ضرب المؤثرات الموسمية" المهم أن لا تكون سنوية"الأربعة أ ، م ، د، ع "

توجد عدة طرق للوصوف إلى الدليل الموسمي، تختلف أساساً في طريقة حساب القيم الاتجاهية:

1/1/2/1/11 طريقة النسبة على المتوسط العام

في هذه الطريقة ناخذ المتوسط العام لكل المواسم على أنه يعكس القيمة الاتجاهية ثم نـتخلص من أثر الاتجاه بقسمة الفعلية لكل موسم على القيمة الاتجاهية، ونأخذ المتوسط الحسابي لبيانات كل موسم للحصول على الدليل الموسمي.

مثال:

قدر الحركة الموسمية للسلسلة التالية التي تمثل مبيعات إحدى المنشآت على أساس ربع سنوي.

	1955	1956	1958	1959
الربع الأول	59	83	106	113
الربع الثاني	76	97	146	147
الربع الثالث	86	91	142	133
الربع الرابع	115	124	188	177

الحل:

1. نحسب المتوسط العام لكافة الأرباح بقسمة المجموع الكلي لقيمة الظاهرة ÷ على الأربـاح بالسلسلة.
2. نحسب متوسط لكل ربع سنة بأخذ مجموع كل ربع (المجموع الأفقي) ثـم نقسـم ÷ عـلى عدد السنوات. وبذلك نصل إلى المتوسط الموسمي.
3. ننسب المتوسط الموسـمي إلى المتوسـط العـام ونضرـب × 100 للحصـول عـلى رقـم قيـاسي موسمي، أو ما يسمى "بالدليل الموسمي".

ويتضمن الجدول التالي الخطوات الحسابية المطلوبة:

					المجموع الموسمي	**المتوسط الموسمي**	**الدليل الموسمي**
الربع الأول	59	83	106	113	455	91	77
الربع الثاني	76	97	146	147	580	116	99
الربع الثالث	86	91	142	133	560	112	95
الربع الرابع	115	124	188	177	755	151	129
					2350		**400**

المتوسط العام $\frac{2350}{20} = 117.5$

"يحتسب المتوسط العام لكافة الأرباع بقسمة المجموع الكلي لقيمة الظاهرة على عدد أرباع السنة".

لحساب المتوسط الموسمي: نقسم مجموع كل موسم على عدد السنوات

$$\frac{445}{5} = 91$$ وهكذا

دليل الموسم:

ثم نقسم المتوسط الموسمي ÷ على المتوسط العام ونضرب × 100 "للحصول على دليل الموسم".

$$\frac{\text{المتوسط الموسمي}}{\text{المتوسط العام}} \times 100 = \frac{91}{117.5} \times 100 = 77.446$$

إن الأرقام القياسية الموسمية أو "دليل الموسم" المحسوب، يوضح أثر الموسم على الظاهرة محل الدراسة -المبيعات مثلاً-.

خلال المواسم الثلاثة الأولى، يكون أثر الموسم بالنقصان، حيث أن المبيعات أقل من المتوسط العام.

ففي الربع الأول مثلاً تمثل المبيعات حوالي ثلاث أرباع المتوسط العام، في حين أنه في الربع الأخير من كل سنة تزيد المبيعات عن المتوسط العام بنحو 29% نتيجة أثر الموسم.

وهذه المعلومات مهمة للمنشأة، فقد تحاول التخفيف من أثر الموسم في الربع الأول بتخفيض الأسعار، أو أنها تستفيد من تلك المعلومات في تخطيط الإنتاج لكل موسم، لتخفيض مصاريف التخزين.

2/1/2/1/11 طريقة النسبة على المتوسط المتحرك

في هذه الطريقة، نحسب أثر الاتجاه العام للظاهرة بطريقة أكثر دقة بدلاً من أخذ المتوسط العام الذي قد يعكس بعض الآثار الدورية والعرضية، ثم نستمر في بقية الخطوات. أي نقسم القيم الفعلية على ÷ القيم الاتجاهية لتخليص القيم الفعلية من أثر الاتجاه، ثم نحسب المتوسط الحسابي، ونضرب في × مائة للحصول على الدليل الموسمي، إلا إذا كان هناك قيم متطرفة، فيفضل استخدام الوسيط بدلاً من المتوسط الحسابي للحصول على رقم قياسي لكل موسم. كما في المثال التالي:

مثال:

احسب الدليل الموسمي باستخدام المتوسط المتحرك على أساس ثلاث فترات:

الموسم	1975	1976	1977	1978	1979
1	42	45	46	48	54
2	55	57	60	64	69
3	24	27	29	31	34
4	19	22	23	24	27

الحل:

الخطوة الأولى: حساب القيم الاتجاهية بأخذ متوسط متحرك لثلاث مواسم ثم نقسم ÷ على (3) للحصول على المتوسط المتحرك الذي يقع بمقابله الموسم الثاني، ثم نسقط الموسم الأول، ونضيف الموسم التالي، وهكذا حسب الجدول التالي:

الموسم	1975	1976	1977	1978	1979
1		40.3	42.6	45.0	49.0
2	40.3	43.0	45.0	47.6	52.3
3	32.6	35.3	37.3	39.6	43.3
4	29.3	31.6	33.3	36.3	-

الخطوة الثانية: قسمة القيمة الفعلية على القيمة الاتجاهية المقابلة لها للتخلص من اثر الاتجاه العام، أي قسمة كل قيمة في الجدول الأول على نظيرتها في الجدول الثاني.

الخطوة الثالثة: أخذ المتوسط الحسابي لكل موسم، أي نجمع الأرقام الخاصة بكل موسم أفقياً، ثم نقسم على عددهم، فالموسم الأول: نأخذ مجموع الصف الأول ونقسم على 4 (أربعة) ونضرب × 100.

الموسم الثاني: مجموع الصف الثاني ونقسم على 5 (خمسة) مع الضرب ×100، وهكذا نحصل على الأرقام الموسمية الخام التي قد نحتاج إلى تعديلها إذا كان مجموعها مختلف عن الرقم (400) نعدل هذه الأرقام للحصول على الأربام الموسمية المعدلة. كما يتضح من الجدول التالي الذي يضم الخطوتين الثانية والثالثة:

الدليل المعدل	الدليل الخام	1979	1978	1977	1976	1975
112	108	1.10	1.06	1.08	1.11	-
139	134	1.32	1.35	1.33	1.33	1.36
80	77	0.97	0.78	0.77	0.76	0.74
69	67	-	0.66	0.69	0.70	0.65
400	386					

وقد تم تعديل الطليل الموسمي الخام بضرب كل منهم في الكسر $\left(\frac{400}{386}\right)$

3/1/2/1/11 طريقة المناسيب المتسلسلة

هذه الطريقة تختصر إحدى الطوات الثلاثة السابقة، حيث نقسم كل مشاهدة على سابقتها، وهكذا نحسب نسبة كل مشاهدة إلى سابقتها مباشرة لكي نتخلص من أثر الاتجاه العام.

في الخطوة الثانية: نحسب المتوسط الحسابي، أو الوسيط، للمناسيب الناتجة، فنحصل على الدليل الموسمي الخام مباشرة، مع ملاحظة أننا قد نحتاج إلى تعديل الأرقام الخام كما سبق وأوضحنا.

مثال:

قدر الحركة الموسمية للظاهرة باستخدام المناسيب المتسلسلة من السلسلة الزمنية التالية:

1959	1958	1957	1956	1955	
7	3	4	6	5	الربع الأول
6	5	5	8	6	الربع الثاني
8	5	7	7	8	الربع الثالث
6	2	3	5	4	الربع الرابع

الحل:

الخطوة الأولى: نحسب المناسيب بقسمة كل مشاهدة على التي تسبقها مباشرة للتخلص من أثر الاتجاه العام فمثلاً 1.20 = $\left(\frac{6}{5}\right)$ القراءة الثانية (6) على القراءة الأولى (5)، ويلي ذلك $1.33=\left(\frac{8}{6}\right)$ القراءة الثالثة على الثانية ... وهكذا لكافة القراءات.

الخطوة الثانية:

حساب لامتوسط الحسابي للمناسيب الخاصة بكل ربع، والضرب في × 100 للحصول على الرقم القياسي الخام. للربع الأول نأخذ الصف الأول ونقسم على عددهم ونضرب × 100 وبالنسبة للربع الثاني، نجمع القيم الواردة بالصف الثاني ثم نقسم على عددهم ونضرب ×100، وهكذا لبقية الفترات.

إلا أن هذه الأرقام يجب أن تعدل حيث أن مجموعها لا يساوي (400) لذا نضرب في النسبة: 400 مقسومة على مجموعهم للحصول على الدليل الموسمي المعدل. نورد الخطوات السابقة في الجدول التالي:

السنة/الفترة	1955	1956	1957	1958	1959	المجموع الموسمي الخام	المتوسط الموسمي المعدل	الدليل الموسمي المعدل
الربع الأول	-	1.50	0.80	1.00	3.50	170	145	
الربع الثاني	1.20	1.33	1.25	1.66	0.85	125	107	
الربع الثالث	1.33	0.87	1.40	1.00	1.33	118	100	
الربع الرابع	0.50	0.71	0.43	0.40	0.75	56	48	
						469	**400**	

استبعاد أثر الموسم:

يمكننا استبعاد أثر الموسم من القيمة الفعلية للظاهرة بقسمة تلك القيم الفعلية على الدليل الموسمي المناظر الذي تم حسابه بإحدى الطرق المذكورة سلفاً، وهكذا نحصل على قيم الظاهرة، كما لو كان لو يكن هناك أثر للموسم.

فمثلاً إذا طلب تخليص الظاهرة في الربع الثاني من سنة 155 في المثال السابق من اثر الموسم نقسم القيمة الفعلية لهذا الموسم أو الربع (6 وحدات) على الدليل الموسمي الخاص بالربع الثاني (107) ونضرب × مائة.

$$\frac{6}{107} \times 100 = 5.60$$

أي أن الظاهرة محل الدراسة كان المتوقع أن تأخذ القيمة (5.60) في المتوسط طبقاً للاتجاه العام، ولكن نظراً لوجود أثر موجب للموسم يجعل الظاهرة، أكبر من المتوسط بمقدار 7%، كما يتضح لنا من الدليل الموسمي، لذا فإن القيمة الفعلية أعلى من المتوسط، بالمثل فإنه يمكن القول بأنه لا يوجد أثر للموسم خلال الربع الثالث، حيث أن الدليل الموسمي يساوي مائة، أي أن الظاهرة لا تختلف في هذا الموسم عن القيمة المتوسطة التي يعكسها الاتجاه العام للظاهرة، وهكذا.

3/1/11 التغيرات الدورية

إن أكثر الطرق انتشاراً لقياس التغيرات الدورية في السلسلة الزمنية، هي طريقة الباقي Residual Method، حيث نبدأ باستبعاد أثر الاتجاه العام ثم أثر الموسم،

ويكون الباقي هو أثر التغيرات الدورية والعرضية وهذه تظهر كنسبة مئوية، لذا تسمى بالمناسيب الدورية Cyclical Relatives.

إن السلسلة الزمنية هي نتيجة حاصل ضرب أربعة مؤثرات، فإذا قسمنا القيم الفعلية على حاصل ضرب القيم الاتجاهية والدليل الموسمي يتبقى لنا أثر التغيرات الدورية والعرضية.

يلاحظ أنه لو كانت السلسلة المتوفرة لدينا سنوية، فإن أثر الموسم لا يظهر في البيانات السنوية، وبالتالي فإنه بعد تخليص المشاهدات من أثر الاتجاه العام يعكس الباقي أثر الدورة الاقتصادية والآثار العرضية، في حين أنه لو كانت السلسلة الزمنية شهرية أو ربع سنوية، لا بد من تخليص القيم الفعلية من أثر الموسم، بالإضافة إلى تخليصها من أثر الاتجاه العام قبل الحصول على المناسيب الدورية.

مثال:

احسب المناسيب الدورية لصادرات دولة ما (z)

T: 1900 1901 1902 1903 1904 1905 1906 1907 1908 1909 1910

Z: 40 39 45 47 46 43 46 57 61 62 63

الحل:

الخطوة الأولى: حساب القيم الاتجاهية باستخدام إحدى الطرق السابقة ولتكن طريقة المربعات الصغرى، كما موضح من الجدول التالي:

T1	X	Z	XZ	X2	$\hat{Z}$	المناسيب الدورية
1900	-5	40	-200	25	37.4	107
1901	-4	39	-156	16	39.9	98
1902	-3	45	-135	9	42.4	106
1903	-2	47	-94	4	44.9	105
1904	--1	46	-46	1	47.4	97
1905	0	43	0	0	49.9	86
1906	1	46	46	1	52.4	88
1907	2	57	114	4	54.9	104
1908	3	61	183	9	57.4	106
1909	4	62	248	16	59.9	103
1910	5	63	315	25	62.4	101
	0	**549**	**275**	**110**		

حيث أننا استخدمنا الانحرافات عن نقطة الوسط:

$$A = \frac{Z}{N} = \frac{549}{11} = 49.9$$

$$B = \frac{XZ}{X^2} = \frac{275}{110} = 2.5$$

(وحدة قياس الزمن السنة، نقطة الأساس سنة 1905)

$$\hat{Z} = 49.9 + 2.5X$$

وللحصول على القيم الاتجاهية، نعوض بقيم المتغير (X) في معادلة الاتجاه العام السابقة، وهو ما تم حسابه في الخانة $\hat{Z}$ ()

فأول قيمة في هذه الخانة هي:

$$\hat{Z}_{1900} = 49.9 + (2.5)\ (-5) = 37.4$$

وهكذا لبقية السنوات.

الخطوة الثانية:

قسمة كل قيمة خطية على القيمة الاتجاهية المناظرة لها × 100 للحصول على المناسيب الدورية التي تمثل نسبة من القيمة "المعتادة" للظاهرة كما تعكسها القيم الاتجاهية.

فمثلاً يمكن القول أن الصادرات في العام الأول للسلسلة كانت أعلى من المعتاد بمقدار 7% تحت تأثير التغيرات الدورية والعرضية، في حين أنه في سنة الأساس (1905) كانت المؤثرات الدورية سالبة بمقدار 14% وهكذا.

وآخر دعوانا، أن الحمد لله رب العالمين

المراجع

المراجع العربية:

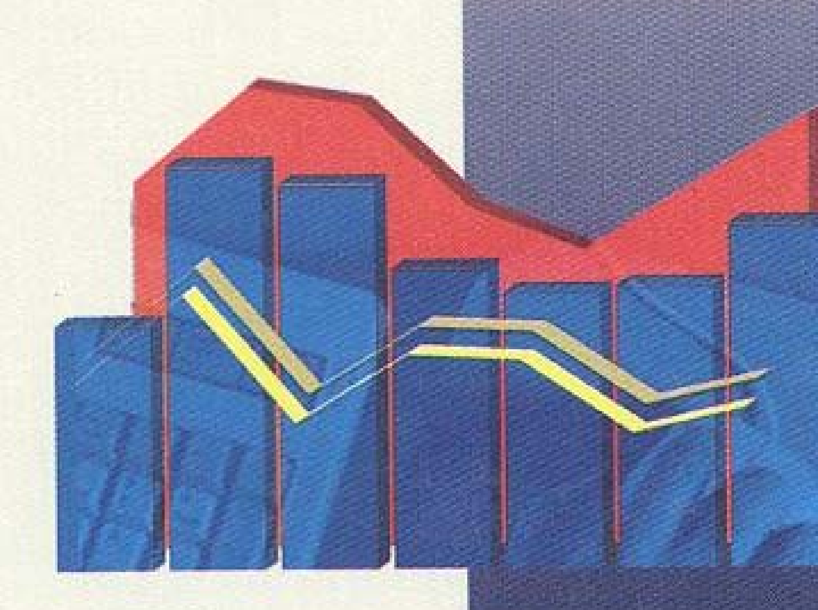
- أبو صالح، محمد صبحي. مقدمة في الإحصاء. جون وايلي. 1983.
- توفيق، ناجي الصالحي، رشيد عبد الرزاق . الإحصاء الهندسي. الطبعة الأولى. جامعة بغداد. 1979.
- الراوي، خاشع محمود. المدخل إلى الإحصاء. جامعة الموصل. 1980.
- خاطر، أحمد، كشك، محمد بهجت. التحليل الإحصائي. القاهرة: المكتب الجامعي الحديث. 1998.
- محمد، علي أبو القاسم. مقدمة في علم الإحصاء التطبيقي. الكويت: المعهد العربي للتخطيط. 1985.
- البدري، وليد، وآخرون. مبادئ الرياضيات للاقتصاديين. الكويت: المعهد العربي للتخطيط. 1985.
- العاني، صبري ديف، إسماعيل، سليم. الطرق الإحصائية. وزارة التعليم العالي والبحث العلمي. 1982.
- منصور، عوض. برمجة بيسك مع تطبيقات. دبي: مكتبة البشائر. 1989.
- الطبولي، أبو القاسم عمر، أبو سدره، فتحي صالح. مبادئ الإحصاء. بنغازي: الدار الجماهيرية للنشر والتوزيع والإعلان. 1993.
- البشير، محمد عثمانن علي، كرم الله. الحاسب الآلي والتطبيقات الإحصائية. السعودية: معهد الإدارة العامة . 1990.

المراجع الأجنبية:

- Huntberger, d.v., Craft, n.j., Billingslet.
- Statistical Inferences for Management of Economics". Boston: Allyn & Bacon Publishing . U.S.A. 1980.
- Van Matic, J.G., Gilbreath, G.H. "Statistics for Business & Economics". Dallas: Business Publication . U.S.A. 1980.
- Spiegel, M; Iheories & Problems of Statistics". N.Y: Mc Graw Hill. 1981.
- Bajpai A.C., Calus I.M. Fairley J.a; "Statistical Methods for Engineers and Scientists". N.J: John Wiley and Sons. 1988.
- Wilfrid J.Dixon, Frank J. Massey Jr. "Introduction to Statistical Analysis" N.Y: Mc Graw Hill. Inc., 1983.
- Freedman D. Lane., "Mathematical Methods In statistics". 1st.W.W. Norton & Company. 1981.
- Crow E.L., Davis F.A., Maxield M.W. "Statistical Mannal". London: Dover, Publication, Inc., 1960.
- Richmond S.B., "Principles of Statistical Analysis". N.J: Ronald Pres Company. 1987.
- Byron S. Gottfried. "Programming with Basic". 3rd ed; N.Y: Mc- Graw-Hill 1986.
- Mandell, West. S.L. "Complete Basic Programming". 1984.
- Summer, M. "Computers: Concepts and uses". N.J: Prentice-Hall. 1985.

الإحصاء التطبيقي

على الحاسوب



دار مجدلاوي للنشر والتوزيع

تليفاكس: ٥٣٤٩٤٩٧ - ٥٣٤٩٤٩٩

ص.ب ١٧٥٨ الجبيهة ١١٩٤١

عمان - الأردن



Dar Majdalawi Pub. & Dis.

Telefax: 5349497 - 5349499

P.O.Box: 1758 Aljubaiha

11941 Amman - Jordan

www.majdalawibooks.com

e-mail: customer@majdalawibooks.com



ISBN 9957-02-188-5